

**MODELOWANIE PREFERENCJI
A RYZYKO '12**

„Studia Ekonomiczne”

**ZESZYTY NAUKOWE
WYDZIAŁOWE
UNIWERSYTETU EKONOMICZNEGO
W KATOWICACH**

MODELOWANIE PREFERENCJI A RYZYKO '12



Katowice 2012

Komitety Redakcyjny

Tadeusz Trzaskalik (redaktor naczelny), Mariusz Żytniewski (sekretarz),
Andrzej Bajdak, Stanisław Stanek, Grażyna Trzpiot

Rada Programowa

Lorenzo Fattorini, Mario Glowik, Miloš Král, Bronisław Micherda,
Zdeněk Mikoláš, Marian Noga, Gwo-Hsiung Tzenga

Redaktor naukowy

Tadeusz Trzaskalik

Recenzenci

Stefan Grzesiak
Donata Kopańska-Bródka
Dorota Kuchta
Maciej Nowak
Wanda Ronka-Chmielowiec
Honorata Sosnowska
Józef Stawicki
Włodzimierz Szkutnik
Grażyna Trzpiot
Tadeusz Tysza

Redaktor

Jadwiga Popławska-Mszyca

© Copyright by Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach 2012

ISBN 978-83-7246-724-9

ISSN 2083-8611

Wszelkie prawa zastrzeżone. Każda reprodukcja lub adaptacja całości bądź części
niniejszej publikacji, niezależnie od zastosowanej techniki reprodukcji,
wymaga pisemnej zgody Wydawcy

WYDAWNICTWO UNIwersytetu Ekonomicznego w Katowicach

ul. 1 Maja 50, 40-287 Katowice, tel. 32 25 77 635, fax 32 25 77 643
www.ue.katowice.pl, e-mail: wydawnictwo@ue.katowice.pl

SPIS TREŚCI

WPROWADZENIE	9
ANALIZA PREFERENCJI	
Marcin Anholcer: ROZWIĄZANIA STABILNE W TRÓJSTRONNYM ZAGADNIENIU PRZYDZIAŁU	17
Summary	22
Jakub Brzostowski, Tomasz Wachowicz: UŻYCIE SKAŁ LINGWISTYCZNYCH DO OPISU UŻYTECZNOŚCI W PROCESIE ANALIZY PREFERENCJI	23
Summary	41
Hanna Bury, Dariusz Wagner: EFEKTYWNE METODY WYZNACZANIA MEDIANY KEMENY'EGO	43
Summary	60
Renata Dudzińska-Baryła: WYBRANE METODY SZACOWANIA FUNKCJI UŻYTECZNOŚCI I FUNKCJI WARTOŚCI	61
Summary	81
Dorota Górecka, Michał Bernard Pietrzak: ZASTOSOWANIE METODY PROMETHEE II W PROCESIE RANKINGOWANIA PROJEKTÓW EUROPEJSKICH W RAMACH REGIONALNEGO PROGRAMU OPERACYJNEGO WOJEWÓDZTWA KUJAWSKO-POMORSKIEGO NA LATA 2007-2013	83
Summary	103
Piotr Jaśkowski, Robert Bucoń, Sławomir Biruk: OCENA PRZYDATNOŚCI UŻYTKOWEJ BUDYNKÓW MIESZKALNYCH Z ZASTOSOWANIEM ROZMYTEGO ROZSZERZENIA AHP	105
Summary	117
Leszek Klukowski: WŁASNOŚCI ESTYMATORÓW RELACJI PORZĄDKU OPARTYCH NA RÓŻNICACH RANG – BADANIE SYMULACYJNE	119
Summary	132

Donata Kopańska-Bródka: WZORZEC MODELU OCZEKIWANEJ UŻYTECZNOŚCI A ROZWÓJ TEORII DECYZJI	133
Summary	148
Joanna Olbryś: WIELOCZYNNIKOWE HYBRYDOWE MODELE MARKET-TIMING POLSKICH FUNDUSZY INWESTYCYJNYCH	149
Summary	161
Aneta Peryga, Honorata Sosnowska: INTUICYJNE I DEDUKCYJNE ROZUMOWANIA W PODEJMOWANIU DECYZJI PRZY UWZGLĘDNIENIU BODŹCÓW FINANSOWYCH I WSPÓŁPRACY GRUPOWEJ	163
Summary	175
Marcin Salamaga: ZASTOSOWANIE ANALIZY CONJOINT DO BADANIA PREFERENCJI MUZYCZNYCH	177
Summary	191
Andrzej M.J. Skulimowski: METODY PROGNOZOWANIA DYNAMICZNYCH RANKINGÓW PRIORYTETÓW TECHNOLOGICZNYCH	193
Summary	211
Jacek Szoltysek, Grażyna Trzpiot: DRZEWA KLASYFIKACYJNE W BADANIU PREFERENCJI KOMUNIKACYJNYCH	213
Summary	230
Tomasz Wachowicz: ANALIZA I WSPOMAGANIE NEGOCJACJI DYSTRYBUCYJNYCH Z ELEMENTAMI ZACHOWAŃ INTEGRACYJNYCH	231
Summary	250

ZARZĄDZANIE PROJEKTAMI

Marek Janczura, Dorota Kuchta: HARMONOGRAMOWANIE REAKTYWNE A RYZYKO W PROJEKCIE	253
Summary	269
Dorota Kuchta: REAKTYWNE HARMONOGRAMOWANIE PROJEKTÓW W PRZYPADKU ROZMYTYCH CZASÓW TRWANIA ZADAŃ I ROZMYTYCH LIMITÓW DOSTĘPNOŚCI ZASOBÓW	271
Summary	284

Aleksandra Niedzielska, Dorota Kuchta: PROBLEM ZAKUPÓW W DUŻYCH PROJEKTACH BADAWCZYCH	285
Summary	296
Anna Ślusarczyk, Dorota Kuchta: ZARZĄDZANIE RYZYKIEM OPÓŹNIENIA PROJEKTÓW REALIZOWANYCH PRZEZ JEDNOSTKI SAMORZĄDOWE O PROSTEJ STRUKTURZE SIECI PROJEKTOWEJ	297
Summary	313
Krzysztof Targiel: METODA DYNAMICZNEGO HARMONOGRAMO- WANIA PORTFELI PROJEKTÓW PRZY OGRANICZONYCH ZASOBACH	315
Summary	324
Katarzyna Tyła, Tomasz Błaszczyk: CZAS I KOSZT REALIZACJI PROJEKTU JAKO KRYTERIA UZUPEŁNIAJĄCE TYPOWY MODEL CAD	325
Summary	336
Łukasz Tync: PODEJŚCIE OBIEKTOWE I WZORZEC OBSERWATOR JAKO WSPARCIE DLA ZARZĄDZANIA RYZYKIEM I KOMUNIKACJĄ W PROJEKCIE	337
Summary	346

RYZIKO W UBEZPIECZENIACH

Stanisław Heilpern: PROCESY RYZYKA Z ZALEŻNYMI WYPŁATAMI	351
Summary	365
Rafał Korzonek: WPLYW RYZYKA DŁUGOWIECZNOŚCI NA UBEZPIECZENIA NA ŻYCIE	367
Summary	375
Anna Nikodem-Słowikowska: ZASTOSOWANIE PROCESU MARKOWA W ZALEŻNYCH GRUPOWYCH UBEZPIECZENIACH NA ŻYCIE	377
Summary	386
Stanisław Wieteska: UBEZPIECZENIA LASÓW OD POŻARÓW W POLSKIEJ STREFIE KLIMATYCZNEJ	387
Summary	401

MISCELLANEA

Sławomir Biruk, Piotr Jaśkowski: OCENA SKUTECZNOŚCI WYBRANYCH METOD REDUKCJI WARIANCJI W BADANIACH SYMULACYJNYCH MODELI SIECIOWYCH PRZEDSIĘWZIĘĆ BUDOWLANYCH	405
Summary	415
Urszula Grzybowska, Marek Karwański: PRZYKŁADY I PORÓWNANIE MODELI MACIERZY MIGRACJI STOSOWANYCH W ANALIZIE RYZYKA KREDYTOWEGO ..	417
Summary	432
Magdalena Rogalska: PROGNOZOWANIE CZASU TRWANIA AWARII KOPAREK KOŁOWYCH Z ZASTOSOWANIEM METOD SIECI NEURONOWYCH, REGRESJI WIELORAKIEJ I ARIMA CZĘŚĆ I	433
Summary	448
Stanisław Walukiewicz: TWIERDZENIE PITAGORASA W NAUKACH SPOŁECZNYCH	449
Summary	474
Aneta Anna Wiktorzak: METODY POMIARU OSIĄGNIĘĆ NA KOLEJNYCH SZCZEBŁACH EDUKACJI	475
Summary	491

WPROWADZENIE

ANALIZA PREFERENCJI

W pracy **Rozwiązania stabilne w trójstronnym zagadnieniu przydziału (M. Anholcer)** autor podaje nowe przykłady zadań, które mają takie rozwiązania.

W pracy **Użycie skal lingwistycznych do opisu użyteczności w procesie analizy preferencji (J. Brzostowski, T. Wachowicz)** przedstawiono propozycję budowy nowej skali lingwistycznej do opisu preferencji decydenta za pomocą termów lingwistycznych, odwzorowywanych na ekwiwalenty numeryczne, i przedyskutowano jej własności.

W pracy **Efektywne metody wyznaczania mediany Kemeny'ego (H. Bury, D. Wagner)** przedstawiono heurystyczną metodę efektywnego wyznaczania tej mediany na podstawie analizy macierzy rozkładu głosów ekspertów, która w istotny sposób ogranicza licznosc zbioru uporządkowań przeszukiwanych przy stosowaniu przeglądu zupełnego.

W pracy **Wybrane metody szacowania funkcji użyteczności i funkcji wartości (R. Dudzińska-Baryła)** dokonano przeglądu wybranych metod określania tych funkcji, przedstawiono także jedną z metod określania funkcji wartości wykorzystywanej do oceny losowych wariantów decyzyjnych na gruncie teorii perspektywy, co zostało zilustrowane rozbudowanym przykładem szacowania funkcji wartości decydenta.

W pracy **Zastosowanie metody PROMETHEE II w procesie rankin-gowania projektów europejskich w ramach Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013 (D. Górecka, M.B. Pietrzak)** przedstawiono aplikację komputerową, której najważniejszą funkcją jest porządkowanie projektów za pomocą metody PROMETHEE II z progiem veta oraz konstruowanie list rankingowych, zgodnie z którymi przyznawane są środki unijne.

W pracy **Ocena przydatności użytkowej budynków mieszkalnych z zastosowaniem rozmytego rozszerzenia AHP** (P. Jaśkowski, R. Bućko, S. Biruk) zaproponowano podejście, które – przy wykorzystaniu nieprecyzyjnej i niepewnej wiedzy pozyskanej od grupy decydentów – umożliwia agregację opinii i ustalenie kompromisu w zakresie istotności kryteriów oceny przydatności użytkowej budynku.

W pracy **Własności estymatorów relacji porządku opartych na różnicach rang – badanie symulacyjne** (L. Klukowski) przedstawiono wyniki badań symulacyjnych w tym zakresie, które wskazują na wysoką efektywność zaproponowanych estymatorów oraz pozwalają określić liczbę porównań.

W pracy **Wzorzec modelu oczekiwanej użyteczności a rozwój teorii decyzji** (D. Kopańska-Bródka) przedstawiono dyskusję dotyczącą głównych nurtów teorii decyzji w kontekście metodologii teorii oczekiwanej użyteczności.

W pracy **Wieloczynnikowe hybrydowe modele market-timing polskich funduszy inwestycyjnych** (J. Olbryś) przedstawiono badanie obejmujące grupę 15 OFI akcji w okresie styczeń 2003-grudzień 2010 przy zastosowaniu metody SUR do estymacji zaproponowanych modeli market-timing w badanej grupie funduszy.

W pracy **Intuicyjne i dedukcyjne rozumowania w podejmowaniu decyzji przy uwzględnieniu bodźców finansowych i współpracy grupowej** (A. Peryga, H. Sosnowska) zawarto opis eksperymentów kontynuujących wcześniejsze badania Sosnowskiej nad intuicyjnym i dedukcyjnym rozumowaniem, ze szczególnym uwzględnieniem koniunkcji i alternatywy zdań.

W pracy **Zastosowanie analizy conjoint do badania preferencji muzycznych** (M. Salamaga) oszacowano użyteczności cząstkowe respondentów, dzięki którym obliczono poziom ważności zmiennych przedstawiających gust, estetykę i kulturę muzyczną badanej grupy melomanów. Wyniki tego rodzaju badań mogą być pomocne w dostosowaniu oferty wydawnictw muzycznych, płytowych, repertuaru sal koncertowych czy programów muzycznych w radiu i telewizji do oczekiwań melomanów.

W pracy **Metody prognozowania dynamicznych rankingów priorytetów technologicznych** (A.M.J. Skulimowski) przedstawiono model, sformułowanie problemu oraz podstawy teoretyczne rozwiązywania pewnej klasy wielokryterialnych problemów decyzyjnych, pojawiających się podczas wyznaczania rankingów kluczowych technologii lub celów strategicznych.

W pracy **Drzewa klasyfikacyjne w badaniu preferencji komunikacyjnych** (J. Szoltysek, G. Trzpiot) rozważono alternatywne metody podróżowania oraz zbadano czynniki, które decydują o preferencjach komunikacyjnych, w tym zamianie istniejącej metody podróżowania na jedną z metod alternatywnych.

W pracy **Analiza i wspomaganie negocjacji dystrybucyjnych z elementami zachowań integracyjnych** (T. Wachowicz) sformułowano autorskie podejście do wspomagania wartościowania ofert w negocjacjach dystrybucyjnych, w których pojawiają się elementy integracyjne między negocjatorami.

ZARZĄDZANIE PROJEKTAMI

W pracy **Harmonogramowanie reaktywne a ryzyko w projekcie** (M. Janczura, D. Kuchta) zaproponowano algorytm, w którym do modelowania niepewności czasów trwania zadań wykorzystano liczby rozmyte; plan projektu aktualizowany był w trakcie jego trwania, biorąc pod uwagę ograniczenia zasobów.

W pracy **Reaktywne harmonogramowanie projektów w przypadku rozmytych czasów trwania zadań i rozmytych limitów dostępności zasobów** (D. Kuchta) zaproponowano częściowo interaktywną metodę wyznaczania początkowego harmonogramu projektu oraz aktualizowania go podczas jego realizacji, uwzględniającą odstęp czasowy między momentem generowania harmonogramu lub jego aktualizacji a momentem planowanego wykonania poszczególnych zadań.

W pracy **Problem zakupów w dużych projektach badawczych** (A. Niedzielska, D. Kuchta), na podstawie badań projektu budowy Wielkiego Zderzacza Hadronów zaproponowano koncepcję wykorzystującą liczby rozmyte, wyboru dostawcy bądź podjęcia decyzji o bliższym sprawdzeniu wybranych dostawców.

W pracy **Zarządzanie ryzykiem opóźnienia projektów realizowanych przez jednostki samorządowe o prostej strukturze sieci projektowej** (A. Ślusarczyk, D. Kuchta) zaproponowano metodę zakładającą, że za każdą ze ścieżek w sieci projektowej odpowiedzialny jest właściciel danej ścieżki, szacując jej planowaną długość oraz oceniając ryzyko, co zwiększa prawdopodobieństwo zakończenia projektu w terminie.

W pracy **Metoda dynamicznego harmonogramowania portfeli projektów przy ograniczonych zasobach** (K. Targiel) przedstawiono zarys metody harmonogramowania portfela projektów, w której potraktowano portfel jako zbiór czynności o niespójnym grafie połączeń, co pozwoliło na wykorzystanie osiągnięć istniejących rozwiązań w zakresie harmonogramowania przy ograniczonych zasobach.

W pracy **Czas i koszt realizacji projektu jako kryteria uzupełniające typowy model CAD** (K. Tyła, T. Błaszczak) przedstawiono model optymalizacyjny wspierający procesy planowania projektu w ujęciu pięciokryterialnym, uzupełniając standardowy dwukryterialny model czasowo-kosztowy o dodatkowe trzy kryteria przestrzenne.

W pracy **Podejście obiektowe i wzorzec obserwator jako wsparcie dla zarządzania ryzykiem i komunikacją w projekcie** (Ł. Tync) wykorzystując analogię pomiędzy zarządzaniem projektami a programowaniem komputerów zapożyczony został pewien zestaw reguł i wzorców stworzonych na potrzeby programowania, które okazały się przydatne w wybranych obszarach zarządzania projektem.

RYZYO W UBEZPIECZENIACH

W pracy **Procesy ryzyka z zależnymi wypłatami** (S. Heilpern) badany jest ciągły proces ryzyka, głównie pod kątem wyznaczania prawdopodobieństwa ruiny w nieskończonym horyzoncie czasu, oraz analizowany jest wpływ intensywności napływu składki na prawdopodobieństwo ruiny.

W pracy **Wpływ ryzyka długowieczności na ubezpieczenia na życie** (R. Korzonek) zaprezentowano zmiany demograficzne oraz rolę ryzyka długowieczności w procesie wyceny składki netto w ubezpieczeniach na życie.

W pracy **Zastosowanie procesu Markowa w zależnych grupowych ubezpieczeniach na życie** (A. Nikodem-Słowikowska) do zobrazowania przyszłego stanu cywilnego mężów i żon zastosowano proces Markowa, pozwala-

jący na wyestymowanie parametrów tego modelu; zbadany także został wpływ zależności między długością życia małżonków na wysokość składki w przypadku renty wdowiej.

W pracy **Ubezpieczenie lasów od pożarów w polskiej strefie klimatycznej (S. Wieteska)** zaprezentowano podstawy kalkulacji stóp składek dla tego rodzaju ubezpieczenia, przedstawiono klasyfikację i przyczyny pożarów, analizę statystyczną kategoryzacji zagrożeń pożarowych, metody szacowania szkód oraz sumy ubezpieczenia.

MISCELLANEA

W pracy **Ocena skuteczności wybranych metod redukcji wariancji w badaniach symulacyjnych modeli sieciowych przedsięwzięć budowlanych (S. Biruk, P. Jaśkowski)** analizowano metody Quasi-Monte Carlo, losowania przeciwnastawnego oraz warstwowego dla modeli sieciowych z czynnościami o czasach realizacji opisanych rozkładami trójkątnymi, stosowanymi w budownictwie do modelowania wpływu zjawisk losowych na parametry organizacji robót.

W pracy **Przykłady i porównanie modeli macierzy migracji stosowanych w analizie ryzyka kredytowego (U. Grzybowska, M. Karwański)** porównano właściwości modeli Markowa uzyskanych na podstawie różnych estymatorów oraz statystycznych longitudinalnych modeli liniowych i wpływ ich zastosowania na wycenę portfela bankowego.

W pracy **Prognozowanie czasu trwania awarii koparek kołowych z zastosowaniem metod sieci neuronowych, regresji wielorakiej i ARIMA część I (M. Rogalska)** przedstawiono metodykę sporządzania arkuszy danych, przekształceń zmiennych oraz doboru modeli pod względem prawidłowości statystycznej ich stosowania, przedstawiając wady i zalety proponowanych metod.

W pracy **Twierdzenie Pitagorasa w naukach społecznych (S. Walkiewicz)**, bazując na zasadzie ortogonalności i równaniu fundamentalnym podano analog twierdzenia Pitagorasa w naukach społecznych, takich jak ekonomia, zarządzanie, socjologia czy nauki polityczne.

W pracy **Metody pomiaru osiągnięć na kolejnych szczeblach edukacji (A. Wiktorzak)** omówiono metody pomiaru/szacowania wartości kapitału ludzkiego oraz zaprezentowano wyniki obszernego eksperymentu, w którym prze-

badano osiągnięcia 156 uczniów/studentów na kolejnych szczeblach edukacji od szkoły podstawowej, przez gimnazjum, liceum, aż po pierwszy rok studiów wyższych i którego zasadniczym celem było porównanie praktycznej użyteczności metody regresji liniowej ze sztucznymi sieciami neuronowymi.

Tadeusz Trzaskalik

ANALIZA PREFERENCJI

Marcin Anholcer

Uniwersytet Ekonomiczny w Poznaniu

ROZWIĄZANIA STABILNE W TRÓJSTRONNYM ZAGADNIENIU PRZYDZIAŁU

Wprowadzenie

W 1962 roku Gale i Shapley w pracy [5], analizując problem przydziału kandydatów do szkół, sformułowali dwustronne zagadnienie przydziału. W zagadnieniu tym występują dwa n -elementowe zbiory: kobiet W i mężczyzn M . Każda kobieta $w \in W$ ma określoną relację silnej preferencji $\succ_w$ na zbiorze M (przy tym zapis $m_1 \succ_w m_2$ oznacza, że kobieta w woli mężczyznę m_1 od mężczyzny m_2), zaś każdy mężczyzna $m \in M$ analogicznie określoną relację silnej preferencji $\succ_m$ na zbiorze W . Przydziałem nazywamy funkcję

$$\mu: W \cup M \rightarrow W \cup M$$

taką, że

$$\mu(w) \in M, \quad \text{dla } w \in W \quad (1)$$

$$\mu(m) \in W, \quad \text{dla } m \in M \quad (2)$$

$$\mu(m) = w \Leftrightarrow \mu(w) = m, \quad \text{dla } w \in W, m \in M \quad (3)$$

Para blokująca przydział μ to para (w, m) , dla której zachodzą warunki

$$m \succ_w \mu(w) \quad (4)$$

$$w \succ_m \mu(m) \quad (5)$$

Gale i Shapley wykazali, że dla każdego układu preferencji istnieje przydział stabilny, tzn. taki, w którym nie występuje żadna para blokująca.

Problem był analizowany i uogólniany na różne sposoby. Dość dokładny przegląd wyników związanych z tym zagadnieniem można znaleźć w przeglądowej pracy Iwamy i Miyazakiego [7].

1. Trójstronne zagadnienie przydziału

Jednym z uogólnień omawianego problemu jest wielostronne zagadnienie przydziału, w którym w miejsce dwóch zbiorów W i M rozpatrujemy większą ich liczbę (powiedzmy k zbiorów), a preferencje są zdefiniowane nie na zbiorze pojedynczych elementów, ale ich uporządkowanych $(k-1)$ -tek. W szczególności rozpatrywane jest trójstronne zagadnienie przydziału ($k=3$), w którym trzeci zbiór oznaczany jest w literaturze np. jako zbiór dzieci C , psów D albo kotów C . W dalszej części pracy posługiwać się będziemy ostatnim z wymienionych oznaczeń, używanym np. w pracy [2]. Trójstronne zagadnienie przydziału możemy wówczas zdefiniować w następujący sposób.

Dane są trzy n -elementowe zbiory: kobiet W , mężczyzn M i kotów C . Każda kobieta $w \in W$ ma określoną relację silnej preferencji $\succ_w$ na zbiorze $M \times C$, każdy mężczyzna $m \in M$ analogicznie określoną relację silnej preferencji $\succ_m$ na zbiorze $W \times C$, a każdy kot $c \in C$ – relację silnej preferencji $\succ_c$ na zbiorze $W \times M$. Przydziałem nazywamy funkcję

$$\mu: W \cup M \cup C \rightarrow (W \times M) \cup (W \times C) \cup (M \times C)$$

taką, że

$$\mu(w) \in M \times C, \quad \text{dla } w \in W \quad (6)$$

$$\mu(m) \in W \times C, \quad \text{dla } m \in M \quad (7)$$

$$\mu(c) \in W \times M, \quad \text{dla } c \in C \quad (8)$$

$$\mu(m) = (m, c) \Leftrightarrow \mu(m) = (w, c) \Leftrightarrow \mu(c) = (w, m) \quad (9)$$

dla $w \in W, m \in M, c \in C$

Trójka blokująca przydział μ to trójka (w, m, c) , dla której zachodzą warunki

$$(m, c) \succ_w \mu(w) \quad (10)$$

$$(w, c) \succ_m \mu(m) \quad (11)$$

$$(w, m) \succ_c \mu(c) \quad (12)$$

W przeciwieństwie do dwustronnego zagadnienia przydziału, zagadnienie trójstronne (ani ogólniej wielostronne) nie musi posiadać przydziału stabilnego. Przykład takiego zagadnienia podany został w pracy [2]. Niech $n=2$, $W = \{w_1, w_2\}$, $M = \{m_1, m_2\}$, $C = \{c_1, c_2\}$, zaś preferencje są postaci

$$(m_2, c_2) \succ_{w_1} (m_2, c_1) \succ_{w_1} (m_1, c_1) \succ_{w_1} (m_1, c_2)$$

$$(m_2, c_2) \succ_{w_2} (m_1, c_1) \succ_{w_2} (m_2, c_1) \succ_{w_2} (m_1, c_2)$$

$$(w_2, c_1) \succ_{m_1} (w_1, c_2) \succ_{m_1} (w_2, c_2) \succ_{m_1} (w_1, c_1)$$

$$(w_2, c_1) \succ_{m_2} (w_1, c_1) \succ_{m_2} (w_2, c_2) \succ_{m_2} (w_1, c_2)$$

$$(w_2, m_1) \succ_{c_1} (w_1, m_2) \succ_{c_1} (w_1, m_1) \succ_{c_1} (w_2, m_2)$$

$$(w_1, m_1) \succ_{c_2} (w_2, m_1) \succ_{c_2} (w_2, m_2) \succ_{c_2} (w_1, m_2)$$

Jak łatwo sprawdzić, przydział $\{(w_1, m_1, c_1), (w_2, m_2, c_2)\}$ blokowany jest przez trójkę (w_1, m_2, c_1) , przydział $\{(w_1, m_2, c_1), (w_2, m_1, c_1)\}$ przez trójkę (w_2, m_2, c_2) , a dwa pozostałe przydziały przez trójkę (w_2, m_1, c_1) .

Jednym z głównych problemów stojących przed badaczami jest podanie warunków koniecznych i wystarczających istnienia przydziału stabilnego. Do tej pory ukazało się kilka prac, w których badano warunki wystarczające. W szczególności analizowano zagadnienia przydziału z cyklicznymi preferencjami. Problem został postawiony przez Knutha w [8], a następnie badany, między innymi przez Borosa i innych ([2], dowód istnienia stabilnych przydziałów dla ściśle cyklicznych, niekoniecznie silnych, relacji preferencji, gdy $n \leq k$), Erikssona i innych ([4], ściśle cykliczne preferencje, $n = 4$, $k = 3$), Huanga ([6], aspekty algorytmiczne i związane ze złożonością problemu) oraz Biró i McDermida ([1], złożoność).

Pewien specyficzny typ preferencji rozpatrywany jest przez Danilova w [3]. Autor definiuje preferencje leksykograficzne, czyli takie, że każdy mężczyzna jest zainteresowany przede wszystkim wyborem kobiety i vice versa (oczywiście każdy ze zbiorów W i M może być zastąpiony przez zbiór C). Niżej przeanalizujemy układy preferencji, w którym preferencje jednej ze grup muszą co prawda pozostać leksykograficzne, jednak nie muszą dotyczyć tylko jednego z pozostałych podzbiorów.

2. Nowy wynik

Głównym rezultatem prezentowanym w niniejszym artykule jest następujące twierdzenie, nawiązujące do wyniku z pracy [3].

Twierdzenie 1

Niech dane będzie trójstronne zagadnienie przydziału. Załóżmy, że spełnione są warunki

1. Każdy mężczyzna $m \in M$ ma relację preferencji $>^*_m$ określoną na zbiorze W , a jego relacja preferencji $>_m$ spełnia warunek

$$w_1 >^*_m w_2 \Rightarrow (w_1, c) >_m (w_2, c), \quad \text{dla } w_1, w_2 \in W, c \in C$$

czyli jest zainteresowany przede wszystkim wyborem kobiety, a wybór kota jest drugorzędny.

2. Każdy kot $c \in C$ ma relację preferencji $>^*_c$ określoną na zbiorze W , a jego relacja preferencji $>_c$ spełnia warunek

$$w_1 >^*_c w_2 \Rightarrow (w_1, m) >_c (w_2, m), \quad \text{dla } w_1, w_2 \in W, m \in M$$

czyli jest zainteresowany przede wszystkim wyborem kobiety, a wybór mężczyzny jest drugorzędny.

3. Zbiór kobiet W można podzielić na dwa podzbiory W_M i W_C , przy czym kobiety $w \in W_M$ są zainteresowane przede wszystkim wyborem mężczyzny, zaś $w \in W_C$ – kota, czyli $w \in W_M$ mają relacje preferencji $\succ_w^M$ zdefiniowane na zbiorze M takie, że

$$m_1 \succ_w^M m_2 \Rightarrow (m_1, c) \succ_w (m_2, c), \quad \text{dla } m_1, m_2 \in M, c \in C$$

zaś $w \in W_C$ relacje preferencji $\succ_w^C$ zdefiniowane na zbiorze C takie, że

$$c_1 \succ_w^C c_2 \Rightarrow (c_1, m) \succ_w (c_2, m), \quad \text{dla } c_1, c_2 \in C, m \in M$$

Wówczas istnieje przydział stabilny.

Dowód

Dowód zaczniemy od konstrukcji przydziału, a następnie wykazemy, że jest on stabilny. Wykorzystamy przy tym kilkakrotnie twierdzenie o istnieniu przydziału stabilnego w dwustronnym zagadnieniu przydziału.

Niech $|W_M| = n_1$, i $|W_C| = n_2$ (oczywiście $n_1 + n_2 = n$). Bazując na relacjach $\succ_m^*$ i $\succ_w^M$ tworzymy μ_1 – częściowy przydział stabilny elementów zbioru W_M i n_1 spośród elementów zbioru M (oznaczymy ich zbiór przez M_1 , a zbiór pozostałych n_2 elementów zbioru M przez M_2). Podobnie, bazując na relacjach $\succ_c^*$ i $\succ_w^C$ tworzymy częściowy przydział stabilny μ_2 wszystkich elementów zbioru W_C i n_2 elementów zbioru C (oznaczymy ich zbiór przez C_2 , a zbiór pozostałych n_1 elementów zbioru C przez C_1). Powstałe w ten sposób pary uzupełniamy do kompletnych trójek w sposób opisany poniżej.

Rozpocznijmy od par (w, c) , gdzie $w \in W_C, c \in C_2$. Dla każdej kobiety $w \in W_C$, po ustaleniu $c \in C_2$, obcięcie relacji $\succ_w$ jest relacją określoną na zbiorze M (oznaczymy ją przez $\succ_w^2$). Przypisujemy tak zdefiniowane relacje parom (w, c) . Jednocześnie zauważmy, że ze względu na uprzednie połączenie $w \in W_C$ i $c \in C_2$ w pary, obcięcie każdej z relacji $\succ_m$ dla $m \in M_2$ definiuje ten sam porządek, co relacja $\succ_m^*$. Dołączamy do każdej z par elementy $m \in M_2$ w taki sposób, aby powstały częściowy przydział μ_2^* par (traktowanych jak pojedyncze elementy) i elementów $m \in M_2$ był stabilny ze względu na relacje $\succ_w^2$ i $\succ_m^*$.

Analogicznie postępujemy w przypadku par (w, m) , gdzie $w \in W_M, m \in M_1$. Dla każdej kobiety $w \in W_M$, po ustaleniu $m \in M_1$, definiujemy relacje $\succ_w^1$ analogicznie, jak w poprzednim przypadku $\succ_w^2$, i przypisujemy je parom (w, m) . Podobnie jak wyżej, dołączamy do każdej z par elementy $c \in C_1$ w taki sposób, aby powstały częściowy przydział μ_1^* par (traktowanych jak pojedyncze elementy) i elementów $c \in C_1$ był stabilny ze względu na relacje $\succ_w^1$ i $\succ_c^*$.

Pokażemy, że w tak zdefiniowanym przydziale μ nie może istnieć trójka blokująca, a więc musi być stabilny.

Założmy, że trójka blokująca istnieje i jest nią (w, m, c) . Założmy ponadto, że $\mu(w) = (m_1, c_1)$, $\mu(m) = (w_2, c_2)$, $\mu(c) = (w_3, m_3)$. Przyjmijmy, że $w \in W_M$ (ze względu na symetrię rozpatrywanego problemu dowód dla $w \in W_C$ może zostać pominięty). Może zajść jeden z dwóch przypadków.

1) $m_1 \neq m$, a więc również $w_2 \neq w$. Oznaczałoby to jednak, że $w_2 \succ_m^* w$ i $m_1 \succ_w^M m$, co byłoby sprzeczne ze stabilnością częściowego przydziału μ_1 ,

2) $m_1 = m$, a więc również $w_2 = w$. Oznaczałoby to jednak, że $c \succ_w^1 c_1$ oraz $(m, w) \succ_c^* (m_3, w_3)$, co byłoby sprzeczne ze stabilnością częściowego przydziału μ_1^* .

Ponieważ w obu przypadkach otrzymaliśmy sprzeczność, para blokująca nie może istnieć. To kończy dowód, gdyż ze względu na symetrię definiowanego zagadnienia analiza przypadku $w \in W_C$ przebiega analogicznie.

Podsumowanie

W pracy zdefiniowana została nowa rodzina trójstronnych zagadnień przydziału, dla których istnieją przydziały stabilne. Rodzina ta jest różna od zdefiniowanych wcześniej, choć przedstawiony rezultat nawiązuje wyników zawartych w pracy Danilova [3].

W dalszych badaniach warto się skoncentrować na próbie uogólnienia otrzymanego wyniku na problemy wielostronne ($k > 3$).

Literatura

1. Biró P., McDermid E. (2010). Three-Sided Stable Matchings with Cyclic Preferences. *Algorithmica*, 58, s. 5-18.
2. Boros E., Gurvich V., Jaslar S., Krasner D. (2004). Stable Matchings in Three-sided Systems with Cyclic Preferences. *Discrete Mathematics*, 289, s. 1-10.
3. Danilov V.I. (2003). Existence of Stable Matchings in Some Three-Sided Systems. *Mathematical Social Sciences*, 46, s. 145-148.
4. Eriksson K., Sjöstrand J., Strimling P. (2006). Three-dimensional Stable Matching with Cyclic Preferences. *Mathematical Social Sciences*, 52, s. 77-87.
5. Gale D., Shapley L.S. College Admissions and the Stability of Marriage. *Am. Math. Mon.*, 69, s. 9-15.
6. Huang C.-C. (2010). Circular Stable Matching and 3-way Kidney Transplant. *Algorithmica*, 58, s. 137-150.

7. Iwama K., Miyazaki S. (2008). A Survey of the Stable Marriage Problem and Its Variants. ICKS '08 Proceedings of the International Conference on Informatics Education and Research for Knowledge-Circulating Society. IEEE Computer Society Washington, DC, USA.
8. Knuth D.E. (1997). Marriages stables et leurs relations avec d'autres problèmes combinatoires (Stable Marriage and its Relation to other Combinatorial Problems). CRM Proceedings and Lecture Notes, Vol. 10. Les Presses de l'Université de Montréal, AMS.

STABLE SOLUTIONS IN THREE-SIDED MATCHING PROBLEM

Summary

In contrast to two-sided matching problem, multi-sided (in particular three-sided) problem not necessarily possesses any stable solution. The author defines new class of problems, for which the existence of such solutions is guaranteed.

Jakub Brzostowski

Politechnika Śląska w Gliwicach

Tomasz Wachowicz

Uniwersytet Ekonomiczny w Katowicach

UŻYCIE SKAL LINGWISTYCZNYCH DO OPISU UŻYTECZNOŚCI W PROCESIE ANALIZY PREFERENCJI*

Wprowadzenie

Jednym z najważniejszych elementów procesu podejmowania decyzji w warunkach wielorakości celów jest identyfikacja i analiza preferencji decydentów. Analiza taka pozwala zdefiniować pewien system oceny wariantów decyzyjnych decydenta i – w konsekwencji – uporządkować warianty decyzyjne od najbardziej do najmniej przez niego preferowanych, identyfikując tym samym wariant najkorzystniejszy. Różnych rekomendacji co do sposobów analizy preferencji dostarczają liczne metody wielokryterialnego podejmowania decyzji MCDM (Multiple Criteria Decision Making). Wśród podejść najczęściej wykorzystywanych do wielokryterialnej analizy preferencji można wyróżnić bazującą na teorii wieloatrybutowej użyteczności MAUT (Multi-Attribute utility Theory) metodę scoringu addytywnego [18], metodę AHP (Analytic Hierarchy Process) [32; 33], metody klasy ELECTRE (ELimination Et Choix Traduisant la REalité [31]) czy PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations) [3]. Metoda scoringu addytywnego syntezyzuje wszystkie kryteria decydenta do jednego kryterium wypadkowego. Zakłada przy tym możliwość określenia jednokryterialnych funkcji oceny, których wartości wyznaczane dla konkretnych poziomów realizacji zapewnianych przez kolejne warianty dają się liniowo agregować za pomocą sumy ważonej do postaci abstrakcyjnego syntetycznego kryterium skalarnego zwanego często użytecznością. Metoda AHP bazuje na dekompozycji wyjścio-

* Praca naukowa finansowana ze środków na naukę w latach 2009-2012 jako projekt badawczy Ministerstwa Nauki i Szkolnictwa Wyższego numer N N111 362337.

wego problemu decyzyjnego na elementarne części składowe i wielokrotnych porównaniach parami tych elementów (np. wariantów czy kryteriów) w celu wyznaczenia ocen cząstkowych. Oceny te, podobnie jak przy scoringu, następnie agregowane są w celu wyznaczenia końcowego wektora ocen opisującego w skali ilorazowej ranking wariantów. Natomiast metody klasy ELECTRE oraz PROMETHEE opierają się na relacji przewyższenia. Dekomponują problem do serii jednokryterialnych porównań parami i wyznaczania syntetycznych wskaźników mierzących sposób przewyższania wariantów w zależności od subiektywnie wyznaczanych progów preferencji, indyferencji czy weta.

Wszystkie wymienione wyżej metody mają swoje wady i zalety. Przykładowo, poza MAUT pozostałe metody nie generują pełnego systemu scoringowego [36], lecz pozwalają jedynie na ocenę predefiniowanych wariantów w zdefiniowanym dyskretnym problemie decyzyjnym. Metody ELECTRE i PROMETHEE przyjmują arbitralnie pewien sposób definiowania preferencji, uwalniając decydenta od szeregu uciążliwych porównań i subiektywnego przypisywania punktów oceny (lub definiowania postaci funkcji oceny), oczekując jednak sprecyzowania wartości pewnych specyficznych parametrów, których interpretacja jest nieintuicyjna, a prawidłowe określenie wymaga sporej wiedzy z zakresu teorii podejmowania decyzji. Natomiast metoda AHP, mimo użycia mocnej skali ilorazowej do opisu porównań wariantów między sobą, pozwala decydentowi na wyrażanie relacji między wariantami za pomocą dziewięciostopniowej skali werbalnej. Taka definicja preferencji wydaje się być bardziej intuicyjna i prostsza dla decydenta. Sam algorytm metody AHP wymaga jednak, aby ocenę werbalną przypisać każdej parze porównywanych elementów składowych problemu. Nawet w prostych problemach decyzyjnych takich par jest zazwyczaj bardzo wiele, dlatego też często pojawia się problem braku spójności zdefiniowanych przez decydenta werbalnie preferencji [21]. W przypadku MAUT, definiowanie preferencji za pomocą skal werbalnych jest kłopotliwe metodologicznie. Syntezowanie kryteriów wyklucza bowiem możliwość stosowania wyrażań werbalnych, gdyż wymaga arytmetycznej agregacji wartości wag i użyteczności, które w takim przypadku zwyczajowo wyrażone są na słabej skali porządkowej. Idea oceny werbalnej nie sprawdzi się również w przypadku metod przewyższenia, te bowiem syntezują predefiniowane funkcje preferencji jednokryterialnych do wypadkowej funkcji preferencji.

Pojawia się zatem pytanie o możliwość praktycznego i skutecznego wykorzystania prostych i intuicyjnie interpretowanych skal werbalnych do definiowania preferencji decydenta w rzeczywistych problemach decyzyjnych. W niniejszej pracy przedstawiamy propozycję autorskiej metody oceny użyteczności za pomocą predefiniowanych skal lingwistycznych, różniącą się od innych podejść zarówno filozofią definiowania takiej oceny, jak również poziomem agregacji problemu, na którym taka definicja następuje. W naszym podejściu wykorzystujemy bowiem sekwencję zbiorów wariantów decyzyjnych (zbiorów referencyjnych) jako punkt odniesienia do wygenerowania globalnego systemu

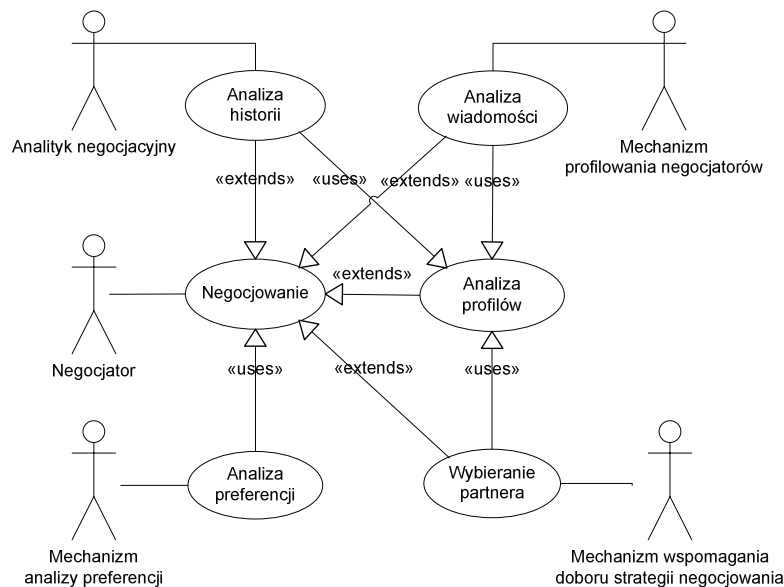
oceny ofert. Zbiór referencyjny składa się z wariantów zdefiniowanych przez decydenta, dla których zakłada się równoważność w sensie preferencji. Tak zdefiniowanemu zbiorowi referencyjnemu decydecnt przypisuje później poziom użyteczności w postaci lingwistycznej. Przyjmujemy bowiem, iż wcześniej decydecnt kalibruje werbalną skalę użyteczności poprzez przypisanie do każdego poziomu tej skali numerycznego ekwiwalentu w postaci oceny użyteczności. Taka procedura pozwala na stosowanie podwójnego opisu skali użyteczności, mianowicie werbalno-numerycznej. Sekwencja zbiorów równoważności z przypisanymi do nich użytecznościami stanowi opis preferencji decydenta. W tak zdefiniowanym systemie referencyjnym (systemie oceny) dowolny wariant decyzyjny oceniany jest w zależności od jego ułożenia względem istniejących zbiorów równoważnych. Im bliżej danego zbioru referencyjnego znajduje się dany wariant tym bliższa jest jego użyteczność użyteczności przypisanej do tego zbioru. Jak więc widać, zaproponowane przez nas podejście nie opiera się ani na syntezie kryteriów, ani na porównywaniu parami wariantów decyzyjnych, z punktu widzenia metodologii omija więc ograniczenia formalne w stosowalności skal werbalnych.

Na całość pracy składają się trzy kolejne rozdziały i zakończenie. W rozdziale 1 przybliżony zostaje kontekst prowadzonych badań. Omówiony zostaje system wspomagania negocjacji NegoManage, na potrzeby którego opracowano przybliżoną wcześniej metodę definiowania preferencji. Opisana zostaje krótko architektura systemu oraz idea zbiorów wariantów równoważnych jako podstawa implementacji skali werbalnej do budowy systemu oceny ofert negocjacyjnych. Rozdział 2 zawiera rozważania dotyczące sposobów definiowania preferencji za pomocą skal ilościowych oraz skal werbalnych, przybliża zagadnienia związane z przejściem między poszczególnymi skalami oraz wskazuje na pewne problemy jakie wiążą się z użyciem skal werbalnych w ocenie preferencji. W rozdziale 3 przedstawiamy szczegółowy opis nowatorskiej skali werbalnej proponowanej do oceny kompletnych wariantów decyzyjnych. Opisujemy podstawy teoretyczne proponowanego podejścia, zagadnienie kalibrowania skali oraz procedurę wyznaczania użyteczności numerycznych na skali zintegrowanej. Pracę kończy podsumowanie, w którym nakreślamy również ramy przyszłych prac związanych z wykorzystaniem oraz koniecznością testowania użyteczności proponowanego podejścia w praktyce negocjacyjnej.

1. System NegoManage

Proponowana w niniejszej pracy koncepcja pomiaru preferencji na podstawie skal lingwistycznych stanowi element składowy większego projektu badawczego, umiejscowionego w kontekście negocjacyjnym, zmierzającego do opracowania i budowy teoretycznego modelu negocjacji dwustronnych, a następnie oprogramowania elektronicznego systemu wspomagania negocjacji

handlowych o dowolnej tematyce i kontekście. System ten, nazwany NegoManage* działa w architekturze rozproszonej [7]. Z punktu widzenia funkcjonalności, zaprojektowany został jako uniwersalna platforma negocjacyjna łatwa do implementacji w postaci podsystemu wspomagającego dowolny serwis typu e-marketplace. System łączy pięciu głównych aktorów: negocjatora, analityka negocjacyjnego, mechanizm analizy preferencji, mechanizm budowania profilu negocjatora oraz jednostkę wspomagającą dobór strategii negocjacyjnej. Wyżej wymienieni aktorzy oddziałują w ramach sześciu ogólnych przypadków użycia (rys. 1). System wspiera działania negocjacyjne aktorów we wszystkich fazach procesu negocjacyjnego, tj. na etapie przygotowawczym (prenegocjacyjnym), podczas negocjacji właściwych (procesu wymiany ofert i argumentacji) oraz w fazie postnegocjacyjnej (w czynnościach analitycznych związanych z analizą Pareto-optymalności uzyskanego wyniku).



Rys. 1. Schemat głównych przypadków użycia systemu NegoManage

W pierwszej fazie negocjacji uczestnicy przygotowują się do rozmów. System NegoManage pomaga w strukturyzacji problemu negocjacyjnego, w ramach której negocjator definiuje kwestie negocjacyjne oraz dopuszczalne poziomy realizacji tych kwestii (zakresy) oraz ich rozdzielczości. Dla tak zdefinio-

* Nazwa jest złożeniem angielskich słów Negotiation i Management określających ideę funkcjonalności samego systemu, który ma pomagać użytkownikowi w sprawnym zarządzaniu swoim procesem negocjacji.

wanego problemu decydent określa swoje preferencje definiując pewne oferty (pakiety kompletne specyfikujące poziomy rozwiązań wszystkich kwestii negocjacyjnych) i grupując je w klasy podobieństwa. Klasy te muszą być później nazwane przez przypisanie im stosownych etykiet jakości. Proces nazywania (oceny) klas wykorzystuje ideę połączonych skal lingwistycznych opisujących użyteczność wyróżnionej grupy ofert (rozdział 4). Negocjator musi również dokonać kalibracji tej skali użyteczności, co pozwala wskazać ekwiwalenty numeryczne wykorzystywanych przez niego termów lingwistycznych i prowadzić dalej proces analityczny (np. związany z postnegocjacyjną poprawą uzyskanych przez negocjujące strony rezultatów).

W fazie drugiej (fazie negocjacji właściwych) system wykorzystuje mechanizm analizy preferencji do oceny ofert negocjacyjnych. Mechanizm ten pozwala na prowadzenie analizy trade-off przy konstruowaniu kolejnych wariantów ustępstw i ocenie ich skali. Wielokrotne operacje trade-off prowadzone dla różnych par kwestii pozwalają bowiem na wygenerowanie zbiorów wariantów równoważnych, którym przypisana zostaje użyteczność werbalno-numeryczna, skalkulowana na podstawie systemu oceny stworzonego w fazie prenegocjacyjnej. Negocjator otrzymuje więc kilka wariantów ustępstw ocenianych przez niego jednakowo, które przedstawić może partnerowi w kolejnych rundach negocjacji jako warianty porozumienia niewymagające kolejnego obniżania jakości diskutowanego kontraktu. Uruchamiany w tej fazie analityk negocjacyjny* jest z kolei odpowiedzialny za analizę wątku negocjacyjnego oraz przygotowanie grafów historii ofert obydwu stron. Pozwala to negocjatorom śledzić postępy procesu negocjacyjnego i analizować jakość wzajemnie czynionych ustępstw. Poza wymianą ofert negocjacyjnych w fazie negocjacji właściwych strony wymieniają również komunikaty tekstowe, wzbogacające proces negocjacyjny o stosowną argumentację. Na podstawie informacji zawartych w komunikatach budowane są profile negocjacyjne stron. Mechanizm budowy profilu analizuje wątek komunikacyjny, aby określić dwa poziomy cech negocjatora, takie jak asertywność czy kooperatywność. Opierając się na zależności między typem wiadomości, która stanowi myśl przewodnią wątku, a typem odpowiedzi na tą myśl, wyliczone zostają poziomy obydwu cech. Kalkulacja przebiega technicznie zgodnie z autorskim algorytmem profilowania [7].

W fazie postnegocjacyjnej system NegoManage uruchamia moduł analityczny odpowiedzialny za analizę efektywności uzyskanego rozwiązania. Analiza ta prowadzona jest w przestrzeni wyników obydwu negocjatorów wobec wyznaczonej uprzednio granicy krańcowo-efektywnej [29]. Mechanizm poszukiwania wariantów lepszych** i optymalnych w sensie Pareto bazuje na idei

* Analityk negocjacyjny jest mechanizmem agentowym, uruchamianym przez system bądź samego negocjatora.

** W sensie zdefiniowanych przez negocjatorów preferencji.

wyznaczania rozwiązania arbitrażowego Gupty-Livne'a (1988) oraz iteracyjnej metody kolejnych przybliżeń granicy Pareto-optymalnej [6]. Taka analiza pozwala uniknąć realizowania przez strony warunków kontraktu nieefektywnego i zabezpiecza negocjatorów przed zjawiskiem opisywanym w teorii negocjacji jako „pozostawienie korzyści na stole negocjacyjnym”, czyli nieskonsumowanie całego tortu negocjacyjnego.

Z całego procesu wspomagania realizowanego przez NegoManage najbardziej istotna, z punktu widzenia wprowadzania lingwistycznej oceny użyteczności, jest idea tworzenia i samo pojęcie zbioru wariantów równoważnych. Zbiór wariantów równoważnych, wygenerowany w procesie analizy preferencji, jest częścią systemu scoringowego pozwalającego na ocenę dowolnej oferty negocjacyjnej w trakcie negocjacji właściwych. Jak wspomniano już wcześniej, użytkownik może definiować taki zbiór wariantów równoważnych stosując wielokrotne operacje trade-off, czyli zaniżając poziomy realizacji pewnych kwestii przy jednoczesnym kompensowaniu ich wyższymi poziomami realizacji kwestii innych. Tenże, zbiór nazwiemy zbiorem referencyjnym, jako, że stanowi on punkt odniesienia do budowy właściwej charakterystyki zbioru wariantów równoważnych (czyli również innych, które cechować się będą tą samą użytecznością). Charakterystyka zbioru wariantów równoważnych zostaje zbudowana w postaci rozkładu prawdopodobieństwa nad zbiorem referencyjnym. Rozkład prawdopodobieństwa jest budowany w oparciu o podejście Kernel Density Estimation [27] w taki sposób, że warianty należące do zbioru referencyjnego mają przypisaną maksymalną wartość prawdopodobieństwa. Natomiast dla innych wariantów prawdopodobieństwo jest niższe i zależy od odległości każdego takiego wariantu od wariantu należącego do zbioru referencyjnego (charakteryzującego się maksymalnym prawdopodobieństwem przynależności). Przypisane w ten sposób prawdopodobieństwo może więc zostać zinterpretowane jako poziom przynależności wariantu do zbioru wariantów równoważnych. System NegoManage stosuje dalej podejście Kernel Density Estimation i buduje wielomodalny rozkład dla każdego zbioru wariantów równoważnych wykorzystując klasteryzację hierarchiczną [14] tak, by w efekcie otrzymać klastry, nad którymi zbudowane zostaną mody. Dopiero taki opis całego systemu referencyjnego można uznać za kompletny system oceny ofert negocjacyjnych.

W przedstawionym wyżej algorytmie wyznaczania systemu oceny ofert negocjacyjnych kluczowy jest odpowiedni dobór sposobu wyrażania oceny generowanych przez negocjatora zbiorów referencyjnych. W kolejnym rozdziale pracy przedyskutujemy pokrótce problematykę definiowania oceny z użyciem skal numerycznych (ilościowych) i werbalnych.

2. Skale numeryczne a skale werbalne

Stosowanie skali werbalnej na jakimkolwiek poziomie oceny problemu decyzyjnego (czy to w definiowaniu istotności kryteriów (wag), czy przy porównywaniu wariantów lub poziomów realizacji kryteriów) uważane jest za prostsze i bardziej intuicyjne, niż operowanie skalą ilościową. Dzieje się tak dlatego, iż w przypadku skal werbalnych decydent używa wyrażen języka naturalnego. Jednakże skali werbalnej brakuje precyzji. Sformułowania typu „mniej”, „dużo mniej”, „średnio” mają subiektywnie różne interpretacje zależne od decydenta, który tych sformułowań używa, przy czym nie jest możliwe również jednoznaczne i proste zdefiniowanie, jak interpretować różnice oceny wyrażone poszczególnymi termami lingwistycznymi. Dzieje się tak dlatego, iż nie istnieje uniwersalny sposób ilościowego (czy statystycznego) określania odległości między dwoma poziomami skali werbalnej. Najtrudniejszym zagadnieniem w praktycznym wykorzystaniu skali werbalnej jest właśnie określenie numerycznej interpretacji wyrażen werbalnych (ekwiwalentów ilościowych) używanych do oznaczania poszczególnych poziomów tejże skali. Z tego punktu widzenia oceny numeryczne trzeba uznać za bardziej precyzyjne od ocen werbalnych. Pozwalają one ponadto (w zależności od typu przyjętej skali numerycznej) na prowadzenie pewnych obliczeń [13], niezbędnych chociażby w późniejszej agregacji cząstkowych ocen wariantów decyzyjnych. Jednakże ludzie często wolą używać wyrażen werbalnych niż numerycznych. Wiele osób niechętnie wyraża opinię w sposób numeryczny, wybierając zamiast operacji na liczbach operacje nienumerycznymi wyrażeniami [10]. Ludzie twierdzą, że „myślą” słowami, nie liczbami i dlatego też lepiej rozumieją znaczenie słów, a nie liczb. Zimmer (1983) twierdzi, że bardziej naturalnym jest myślenie i mówienie o niepewności za pomocą wyrażen werbalnych oraz że ludzie potrafią lepiej stosować zasady języka niż zasady prawdopodobieństwa. Nienumeryczne frazy lepiej oddają właściwą dla opinii niepewność [9]. Oczywiście należy mieć jednocześnie świadomość, iż takie nieostre definiowanie ocen (czy preferencji) w postaci termów lingwistycznych może skutkować nie tylko brakiem precyzji finalnego systemu oceny, ale także pewnym brakiem jego spójności. Przyczyna tego może być dwojaka. Z jednej strony błędy mogą być popełniane przez samych decydentów i dotyczyć jakiejś nieuzasadnionej interpretacji ocen słownych [34]*, z drugiej jednak mogą być związane z przyjętą metodologią przetwarzania i obróbki takich ocen [25; 21].

* Ciekawą i wyczerpującą dyskusję na ten temat znaleźć można w pracy Tyszki i Zalaśkiewicza (2001).

Numeryczna interpretacja wyrażen werbalnych stanowi problem badawczy od trzech dekad. W wielu eksperymentach badacze próbowali ustalić jak ludzie interpretują wyrażenie typu: „możliwie”, „prawdopodobnie”, „prawie na pewno” [2; 4; 12; 28]. Literatura pokazuje, że ludzie różnie przypisują numeryczne wartości wyrażeniom werbalnym. Na przykład, jeśli wiadomość zawierająca wyrażenie werbalne zostaje przesłana od jednego aktora do drugiego, odbiorca może mieć (i zazwyczaj ma) inną interpretację tego wyrażenia niż miał nadawca [12]. To samo wyrażenie może mieścić się w bardzo szerokim zakresie interpretacyjnym różnych decydentów, a co za tym idzie, skutkować znacznym rozrzutem numerycznej interpretacji tego wyrażenia werbalnego. Dlatego też zmierza się do stworzenia pewnych kanonów interpretacyjnych wyrażen werbalnych, pozwalających na powiązanie oceny werbalnej z jej ekwiwalentem numerycznym i uzmysłowieniu decydentowi konsekwencji jakie te oceny – mogą mieć na późniejszy proces analityczno obliczeniowy. Pewne standardy definiowania ekwiwalentów numerycznych dla termów lingwistycznych proponowane są oddzielnie dla problemów określania wag kryteriów [16], czy porównania wariantów między sobą i dotyczą zarówno precyzyjnego określania ekwiwalentów w postaci wartości skalarnych, jak i nieostrej definicji w postaci liczb przedziałowych lub rozmytych [1]. Przykład różnych projekcji opisu lingwistycznego na rozmyte ekwiwalenty numeryczne (w tym przypadku na liczby trójkątne) zaprezentowano w tabeli 1.

Tabela 1

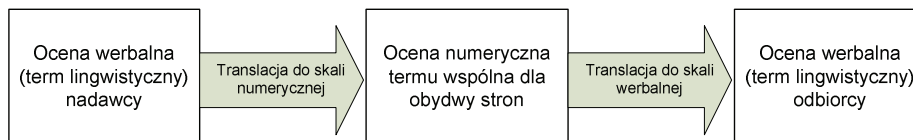
Cechy lingwistyczne wykorzystywane do określania rankingu wariantów decyzyjnych

Cecha lingwistyczna	Ekwiwalent numeryczny cechy (trójkątna liczba rozmyta)	
	Według Ou i Chou (2009)	Według Chen (2000)
Bardzo słaby (BS)	(0, 0, 10)	(0, 0, 1)
Słaby (S)	(0, 10, 30)	(0, 1, 3)
Średnio słaby (ŚS)	(10, 30, 50)	(1, 3, 5)
Dostateczny (D)	(30, 50, 70)	(3, 5, 7)
Średnio dobry (ŚD)	(50, 70, 90)	(5, 7, 9)
Dobry (D)	(70, 90, 100)	(7, 9, 10)
Bardzo dobry (BD)	(90, 100, 100)	(9, 10, 10)

Źródło: Opracowanie własne na podstawie [30].

Alternatywnie, nadawca oraz odbiorca przekazu mogą mieć swoje własne interpretacje werbalne standardowego przekazu numerycznego. Definicja nadawcy odwzorowywana jest w ustandaryzowanej skali numerycznej i przekazywana odbiorcy, który interpretuje ją zgodnie ze swoją skalą. Mimo różnic

opisu werbalnego, istnieje pewność co do istoty zagadnienia, które terminy te opisują po obydwu stronach toczonej dyskusji. Schemat takiego kodowania przedstawiono na rys. 2.



Rys. 2. Translacja wyrażenia werbalnego po stronie nadawcy i przekaz kanałem transmisji wyrażenia numerycznego do odbiorcy

Ze względu na różnorodność interpretacji wyrażen werbalnych analizuje się czynniki wpływające na tę interpretację. Jednym z takich czynników jest doświadczenie decydenta. Trzy prace w dziedzinie podejmowania decyzji w medycynie dają jednak sprzeczne rezultaty. Nakao i Axelrod (1983) odkryli zjawisko większego konsensusu pomiędzy lekarzami niż pomiędzy nieeksper-tami (np. pacjentami) przy definiowaniu numerycznego znaczenia wyrażen werbalnych. Jednakże zespół Konga [20] oraz Timmermans (1994) odkryli później brak wpływu doświadczenia na ową interpretację. Drugim czynnikiem wpływającym na poziom różnorodności interpretacji jest porządek, w jakim wyrażenia werbalne są prezentowane. Hamm (1991) odkrył wyższą różnorodność w przypisywaniu wartości numerycznych do ocen werbalnych, gdy wyrażenia były prezentowane w porządku losowym. Zespół naukowy Clarke'a [10] otrzymał podobne rezultaty. Trzecim czynnikiem, który może mieć wpływ na różnorodność interpretacji wyrażen werbalnych jest kontekst definiowania oceny, tj. sytuacja prezentowania ocenianego wyrażenia w sposób wyizolowany bądź też jako części większego tekstu. Wyniki badań opisanych w literaturze są sprzeczne również i w tej kwestii. Mapes (1979) oraz Bryant i Norman (1980) argumentują, iż numeryczne interpretacje wyrażen werbalnych zależą od kontekstu. Beyth-Marom (1982) oraz Brun i Teigen (1988) potwierdzają, że oceny są bardziej różnorodne gdy wyrażenie pojawia się w kontekście, niż gdy jest oceniane poza kontekstem. Tymczasem Bass (1974) oraz Timmermans (1994) odkryli brak wpływu kontekstu na różnorodność interpretacji. Wydaje się zatem, iż nie ma jednoznacznej relacji pomiędzy różnorodnością interpretacji a kontekstem. Być może, jak to zauważyli Budescu i Wallsten (1985), po prostu sam kontekst nie jest dobrze zdefiniowaną jednowymiarową koncepcją.

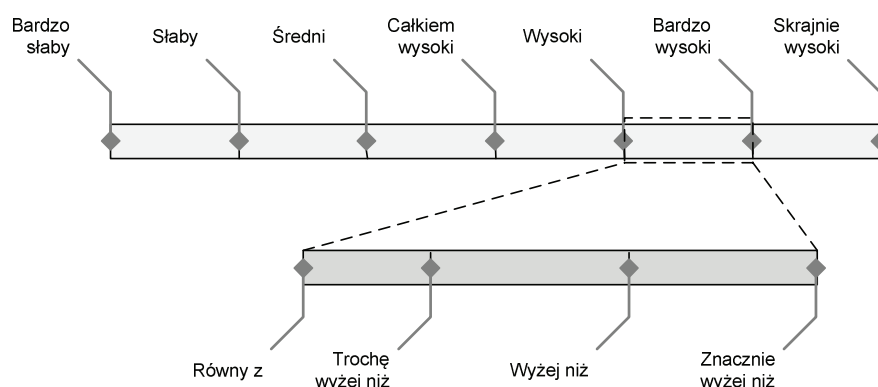
W celu rozwiązania problemu różnych interpretacji wyrażen werbalnych przez dwóch różnych aktorów, w kolejnym rozdziale pracy proponujemy nową metodę definiowania preferencji wykorzystującą podwójną skalę werbalną. Zdefiniowane w ten sposób preferencje będą stosownie kalibrowane

i dekodowane na unormowane ekwiwalenty numeryczne, pozwalające na prowadzenie ewentualnej analizy symetrycznej problemu z punktu widzenia obydwu aktorów jednocześnie, co ma swe uzasadnienie w kontekście negocjacji dwustronnych, gdzie zaimplementowane zostało proponowane przez nas rozwiązanie.

3. Propozycja nowej, zintegrowanej skali werbalnej

Aby nie utracić intuicyjności budowanej skali werbalnej, liczba poziomów tej skali musi być ograniczona. Badania z zakresu psychologii decyzji sugerują użycie maksymalnie siedmiu termów skali werbalnej. Jednakże liczba siedmiu punktów skali redukuje elastyczność oceny wariantów. Interpretowalne odległości między dwoma sąsiednimi poziomami skali mogą być duże, a w sytuacji siedmiu poziomów realizacji zwykle pozostaną niezdefiniowane lub źle uchwycone [21]. W celu uzyskania wyższej precyzji liczba termów powinna być zwiększona. Z drugiej jednak strony, zwiększenie liczby termów zwykłej skali lingwistycznej może doprowadzić do utraty jej intuicyjności, gdyż decydent może przestać odróżniać niektóre z poziomów skali (ich zapis słowny oznaczać będzie dla niego to samo).

W celu rozwiązania powyższego problemu w proponowanym przez nas podejściu stosujemy dwie skale. Pierwsza skala ma na celu jedynie zgrubną ocenę wariantu. Natomiast skala druga ma na celu doprecyzowanie oceny pierwszej, poprzez określenie umiejscowienia wariantu pomiędzy dwoma sąsiednimi termami skali pierwszej. Innymi słowy, w kroku pierwszym oceny wybranego wariantu decyzyjnego decydent określa poziom pierwszej skali, który najlepiej opisuje jakość tego wariantu (definiuje poziom jakości, który jest zapewniany przez ten wariant). W drugim kroku zaś wybiera on poziom umiejscowienia tego wariantu pomiędzy dwoma następującymi po sobie poziomami skali pierwszej (tj. między poziomem, który oceniany wariant zapewnia, a poziomem następnym, którego już nie osiąga), stosując do tego zadania bardziej szczegółową skalę drugą. Druga skala rozbija zakres pomiędzy dwoma sąsiednimi termami skali pierwszej na podpoziomy. Na przykład, wariant może przewyższać poziom bazowy (np. „Wysoki”) skali pierwszej, lecz nie sięgać poziomowi „Bardzo Wysoki”. Druga skala jest stosowana właśnie do określenia stopnia przewyższenia przez oceniany wariant bazowego poziomu „Wysoki”. Wariant może bowiem przewyższać poziom „Wysoki” w niewielkim stopniu i wtedy ostateczny poziom przypisany do wariantu wynosi „Trochę Wyżej niż Wysoki” lub bardzo, a wówczas jego definicja werbalna wyniesie „Znacznie wyżej niż Wysoki”. Ideę takiej skali pomiaru zaprezentowano na rys. 3.



Rys. 3. Ilustracja podwójnej, zintegrowanej skali werbalnej do oceny zgrubnej i szczegółowej wariantu decyzyjnego

Jak pokazano na rys. 3 rozbiecie odległości pomiędzy poziomami skali pierwszej poprzez użycie drugiej skali lingwistycznej pozwala na zwiększenie precyzji oceny wariantu bez utraty intuicyjności oceny. Zakładając, że liczba poziomów obydwu skal jest niska (na przykład nie większa niż 7) pozwala nam to utrzymać intuicyjność całej analizy preferencji niezbędnej do zbudowania pełnego systemu oceny wariantów decyzyjnych. Ponadto, dzięki stosowaniu podwójnej, zintegrowanej skali uzyskujemy wyższą precyzję oceny. Przyjmując, że obydwie skale mają siedem poziomów, fuzja skal daje ostatecznie 37 możliwych poziomów oceny na skali zintegrowanej.

3.1. Kalibracja skali werbalnej

Kalibracja skali lingwistycznej polega na przypisaniu wartości numerycznych (interpretowanych jako użyteczności) do wszystkich poziomów tejże skali. Kalibrując przedstawioną w poprzedniej części pracy pierwszą skalę lingwistyczną proponujemy wykorzystać skalę numeryczną użyteczność najczęściej stosowaną w innych metodach wielokryterialnej oceny, jak na przykład w algorytmach MAUT, gdzie użyteczność definiuje się zazwyczaj w zakresie $[0; 100]$ (lub ekwiwalentnie $[0; 1]$). Taka definicja ocen numerycznych zachowuje pewną intuicyjną interpretację w której najgorzej oceniony wariant otrzymuje scoring równy 0 (zerowa jakość), a najlepiej oceniony – scoring 100 (pełna, 100% jakość). Przykładową kalibrację dwustopniowej skali lingwistycznej przedstawiono w tabeli 2.

Tabela 2

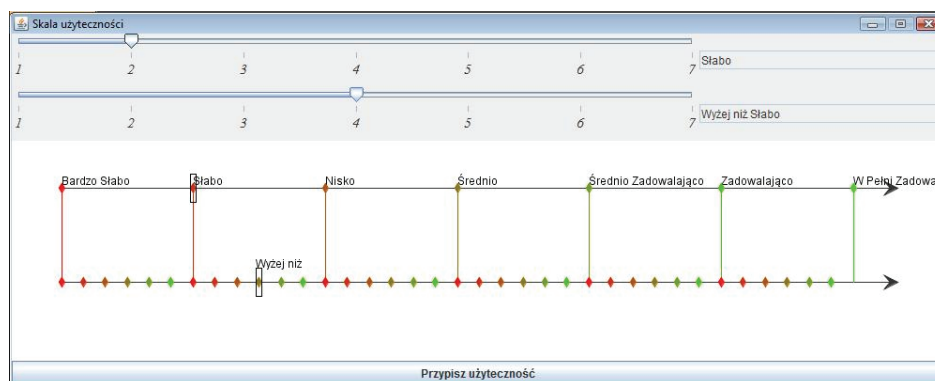
Przykładowa kalibracja podwójnej skali lingwistycznej

Wartość lingwistyczna (Skala pierwsza)	Ekwiwalent numeryczny (użyteczność)
Bardzo Słabo	0
Słabo	15
Nisko	30
Średnio	40
Średnio Zadowolająco	55
Zadowolająco	70
W pełni zadowolająco	100
Wartość lingwistyczna (Skala druga)	Ekwiwalent numeryczny (użyteczność)
Równy z	0
Nieznacznie Wyżej niż	10
Trochę Wyżej niż	20
Wyżej niż	30
Znacznie Wyżej niż	40
Dużo Wyżej niż	50
O Wiele Wyżej niż	60

Z przykładu zapisanego w tabeli 2 wynika, iż użytkownik przypisał do wszystkich termów lingwistycznych pierwszej skali oceny z zakresu 0 do 100, a do termów lingwistycznych drugiej skali wartości numeryczne od 0 do 60. Na pierwszej skali użytkownik ogólnie definiuje poziom użyteczności dla ocenianego wariantu. Poziomy drugiej skali zostają zestandaryzowane w taki sposób, że pierwszy poziom skali drugiej pokrywa się z poziomem skali pierwszej (pierwszy poziom drugiej skali werbalnej definiowany jest jako „równy z”, konsekwentnie więc musi on odpowiadać jakości dokładnie podstawowemu stopniowi skali pierwszej) oraz ostatni poziom skali drugiej pokrywa się z następnym poziomem skali pierwszej (taka interpretacja przypisana jest poziomowi „o wiele wyżej niż”). Zakładając, iż operujemy użytecznością kardynalną różnice w ocenie kolejnych poziomów skali drugiej wyznaczają przyrosty użyteczności na skali pierwszej. W tak zdefiniowanej podwójnej skali werbalnej dysponować będziemy ostatecznie 37 poziomami oceny dostępnymi decydentowi.

Na rys. 4 przedstawiono przykładowy interfejs definicji oceny wariantu w zintegrowanej skali werbalnej, udostępniany przez system NegoManage. Podczas oceny wariantu użytkownik systemu dobiera za pomocą suwaków poziom ze skali pierwszej, a następnie ze skali drugiej doprecyzowującej (jak

widać, na rysunku dobrano ze skali pierwszej poziom „Słabo”). Tak zdefiniowane preferencje werbalne muszą zostać przekodowane na ekwiwalenty numeryczne zgodnie z definicją użyteczności poszczególnych poziomów skali pierwszej i drugiej.



Rys. 4. Ocena wariantów decyzyjnych na zintegrowanej skali werbalnej w systemie Nego-Manage.

3.2. Użyteczności numeryczne skali zintegrowanej

Ostatnim zadaniem związanym z opracowaniem pełnej skali zintegrowanej jest określenie wartości numerycznych odpowiadających wszystkim 37 termom lingwistycznym. Wartości skali numerycznej zostają wyliczone w następujący sposób. Dla poziomów zintegrowanej skali użyteczności pokrywających się z poziomami skali pierwszej wartości numeryczne użyteczności są pobierane z wartości numerycznych pierwszej skali (oznaczonych przez fS_i , gdzie $i \in \{1, \dots, 7\}$ i oznacza numer kolejnego poziomu skali pierwszej).

Dla poziomów pośrednich (poziomów pomiędzy poziomami skali pierwszej) generowanych przy pomocy skali drugiej wartości numeryczne zostają przypisane proporcjonalnie do wartości skali drugiej (wartości skali drugiej oznaczamy z kolei przez sS_j , gdzie $j \in \{1, \dots, 7\}$ i oznacza numer kolejnego poziomu skali drugiej). Użyteczność poziomów pośrednich skali zintegrowanej wyznaczymy zatem z formuły

$$numericUtility(i, j) = \begin{cases} fS_i + (fS_{i+1} - fS_i) \frac{(sS_j - sS_1)}{(sS_7 - sS_1)} & \text{dla } i < 7 \quad j < 7 \\ fS_{i+1} & \text{dla } i < 7 \quad j = 7 \\ fS_i & \text{dla } i = 7 \end{cases} \quad (*)$$

gdzie:

$fS_i + (fS_{i+1} - fS_i) \frac{(sS_j - sS_1)}{(sS_7 - sS_1)}$ jest pośrednim stopniem użyteczności wypadkowej położonej pomiędzy dwoma sąsiednimi poziomami skali pierwszej (fS_i oraz fS_{i+1}), $\frac{(sS_j - sS_1)}{(sS_7 - sS_1)}$ odpowiada wartości poziomu drugiej skali przeskalowanej do przedziału $[0,1]$.

Jeśli $i = 7$ użyteczność wypadkowa $numericUtility(i, j)$ pokrywa się z poziomem pierwszej skali (fS_7), gdyż mamy do czynienia z najwyższą oceną dostępną w utworzonej skali zintegrowanej. Jeśli $j = 7$ oraz $i < 7$ użyteczność wypadkowa $numericUtility(i, j)$ pokrywa się z następnym poziomem skali pierwszej (fS_{i+1}).

W tabeli 3 ilustrujemy pełną skalę zintegrowaną składającą się z 37 ostatecznych termów lingwistycznych wraz z ich ekwiwalentami numerycznymi wyznaczonymi zgodnie z formułą (*). W pierwszej kolumnie tabeli zaprezentowano poziomy skali pierwszej, w drugiej kolumnie uszczegółowienia każdego termu lingwistycznego skali pierwszej termami ze skali drugiej, w kolumnie trzeciej dokonano złożenia, uzyskując tym samym termy skali zintegrowanej. Czwarta kolumna zawiera wartości numeryczne użyteczności odpowiadające termom skali zintegrowanej, wyznaczone zgodnie z definicją decydenta (tabela 2).

Tak zdefiniowaną, skalibrowaną skalę z wyznaczonymi ekwiwalentami numerycznymi w postaci użyteczności kardynalnych wykorzystać można do oceny wariantów decyzyjnych, jak również do dalszych, bardziej zaawansowanych analiz. Przykładowo w kontekście negocyjnym systemy oceny ofert obydwu negocjatorów wyznaczone na podstawie zintegrowanej skali werbalnej nadają się do analizy symetrycznej jakości osiągniętego kompromisu i jego analitycznej poprawy.

Tabela 3

Zintegrowana skala lingwistyczna wraz z ekwiwalentną użytecznością numeryczną

Poziomy pierwszej skali	Poziomy drugiej skali	Nazwa terminu lingwistycznego w skali wypadkowej	Użyteczność numeryczna
1	2	3	4
Bardzo Słabo	Równy z	Równy z Bardzo Słabo	0
	Nieznacznie wyżej niż	Nieznacznie wyżej niż Bardzo Słabo	2,5
	Trochę wyżej niż	Trochę wyżej niż Bardzo Słabo	5
	Wyżej niż	Wyżej niż Bardzo Słabo	7,5
	Znacznie wyżej niż	Znacznie wyżej niż Bardzo Słabo	10
	Dużo wyżej niż	Dużo wyżej niż Bardzo Słabo	12,5
	O wiele wyżej niż	O wiele wyżej niż Bardzo Słabo	15
Słabo	Równy z	Równy z Słabo	15
	Nieznacznie wyżej niż	Nieznacznie wyżej niż Słabo	17,5
	Trochę wyżej niż	Trochę wyżej niż Słabo	20
	Wyżej niż	Wyżej niż Słabo	22,5
	Znacznie wyżej niż	Znacznie wyżej niż Słabo	25
	Dużo wyżej niż	Dużo wyżej niż Słabo	27,5
	O wiele wyżej niż	O wiele wyżej niż Słabo	30
Nisko	Równy z	Równy z Nisko	30
	Nieznacznie wyżej niż	Nieznacznie wyżej niż Nisko	31,67
	Trochę wyżej niż	Trochę wyżej niż Nisko	33,34
	Wyżej niż	Wyżej niż Nisko	35
	Znacznie wyżej niż	Znacznie wyżej niż Nisko	36,68
	Dużo wyżej niż	Dużo wyżej niż Nisko	38,35
	O wiele wyżej niż	O wiele wyżej niż Nisko	40
Średnio	Równy z	Równy z Średnio	40
	Nieznacznie wyżej niż	Nieznacznie wyżej niż Średnio	42,5
	Trochę wyżej niż	Trochę wyżej niż Średnio	45
	Wyżej niż	Wyżej niż Średnio	47,5
	Znacznie wyżej niż	Znacznie wyżej niż Średnio	50
	Dużo wyżej niż	Dużo wyżej niż Średnio	52,5
	O wiele wyżej niż	O wiele wyżej niż Średnio	55
Średnio Zadowolająco	Równy z	Równy z Średnio Zadowolająco	55
	Nieznacznie wyżej niż	Nieznacznie wyżej niż Średnio Zadowolająco	57,5
	Trochę wyżej niż	Trochę wyżej niż Średnio Zadowolająco	60
	Wyżej niż	Wyżej niż Średnio Zadowolająco	62,5
	Znacznie wyżej niż	Znacznie wyżej niż Średnio Zadowolająco	65

cd. tabeli 3

1	2	3	4
	Dużo wyżej niż	Dużo wyżej niż Średnio Zadowolająco	67,5
	O wiele wyżej niż	O wiele wyżej niż Średnio Zadowolająco	70
Zadowolająco	Równy z	Równy z Zadowolająco	70
	Nieznacznie wyżej niż	Nieznacznie wyżej niż Zadowolająco	75
	Trochę wyżej niż	Trochę wyżej niż Zadowolająco	80
	Wyżej niż	Wyżej niż Zadowolająco	85
	Znacznie wyżej niż	Znacznie wyżej niż Zadowolająco	90
	Dużo wyżej niż	Dużo wyżej niż Zadowolająco	95
	O wiele wyżej niż	O wiele wyżej niż Zadowolająco	100
W pełni zadowolająco	Równy z	Równy z W pełni zadowolająco	100

Podsumowanie

W niniejszej pracy zaproponowaliśmy sposób konstrukcji nowej skali werbalnej, pozwalającej na ocenę wariantu decyzyjnego w sposób lingwistyczny. Problem niskiej precyzji skali werbalnej, powiązany z brakiem realnych możliwości (interpretacyjnych) wprowadzania większej liczby terminów lingwistycznych, został rozwiązany poprzez wprowadzenie integracji dwóch skal werbalnych, skali pierwszej, pozwalającej na zgrubną ocenę wariantu oraz skali drugiej, doprecyzowującej ocenę skali pierwszej. Poprzez integrację dwóch skal werbalnych otrzymano możliwość oceny wariantu na 37 stopniowej skali werbalnej. W pracy przedyskutowano wady i zalety skali werbalnej jak również przedstawiono mechanizm odwzorowania skali werbalnej w skalę numeryczną po uprzednim odwzorowaniu dwóch skal składowych w skale numeryczne. Mechanizm ten zaimplementowany został w systemie wspomagania negocjacji NegoManage do budowy systemu oceny ofert negocjacyjnych, a jego uniwersalność pozwala na wykorzystanie tak zdefiniowanych ocen wariantów również w symetrycznej, teoriogrowej analizie całego procesu negocjacyjnego, stwarzając możliwości prowadzenia skutecznej mediacji czy arbitrażu. Nasza propozycja, jak każde nowo opracowane konceptualnie narzędzie, wymaga testów na użytkownikach finalnych. Poprawność metodologiczna bowiem nie jest wystarczająca do uznania narzędzia za użyteczne. Kolejnym etapem planowanych przez nas prac badawczych będzie dokonanie pogłębionej analizy użyteczności i akceptowalności proponowanej skali werbalnej przez negocjatorów, uczestników negocjacji elektronicznych prowadzonych za pomocą systemu NegoManage. Podczas tychże testów badane będą zalety oraz wady zapropo-

nowanej metody, co pozwoli na ewentualne dostosowanie jej do wymogów użytkownika końcowego, przy zachowaniu metodologicznej zgodności i poprawności.

Literatura

1. Bass B.M., Cascio W.F., and O'Conner E.J. (1974). Magnitude Estimation of Expressions of Frequency and Amount. *Journal of Applied Psychology*, 59, s. 313-320.
2. Beyth-Marom R. (1982). How Probable is Probable? Numerical Translation of Verbal Probability Expressions. *Journal of Forecasting*, 1, s. 257-269.
3. Brans J.P., Mareschal B. and Vincke Ph. (1984). PROMETHEE: A New Family of Outranking Methods in Multicriteria Analysis. In: *Operational Research '84*. Ed. J.P. Brans. Elsevier Science Publishers B.V., North-Holland, s. 408-421.
4. Brun W., Teigen K.H. (1988). Verbal Probabilities: Ambiguous, Context-Dependent, or Both? *Organizational Behavior and Human Decision Processes*, 41, s. 390-404.
5. Bryant G.D., and Norman G.R. (1980). Expressions of Probability: Words and Numbers. *The New England Journal of Medicine*, 302, s. 411.
6. Brzostowski J. (2011). Poprawa rozwiązania negocjacyjnego w systemie Nego-Manage poprzez zastosowanie rozwiązania przetargowego. XXX Konferencja Naukowa Metody i Zasotswania Badań Operacyjnych. Trzebieszowice 16-18 października.
7. Brzostowski J., Wachowicz T. (2011). Budowa systemu reputacyjnego z taksonomią aktów mowy do wspomagania działania systemu negocjacji elektronicznych. W: *Wiedza i komunikacja w innowacyjnych organizacjach. Komunikacja elektroniczna*. Red. M. Pańkowska. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice, s. 30-55.
8. Brzostowski J., Wachowicz T. (2009). Conceptual Model of eNS For Supporting Preference Elicitation And Counterpart Analysis, The Joint Conference of the INFORMS Section on Group Decision and Negotiation, the EURO Working Group on Decision and Negotiation Support, and the EURO Working Group on Decision Support Systems, Toronto.
9. Budescu, D.V., and Wallsten, T.S. (1985). Consistency in Interpretation of Probabilistic Phrases. *Organizational Behavior and Human Decision Processes*, 36, s. 391-405.
10. Clarke V.A., Ruffin C.L. Hill D.J., and Beaman A.L. (1992). Ratings of Orally Presented Verbal Expressions of Probability by a Heterogeneous Sample. *Journal of Applied Social Psychology*, 22, s. 638-656.
11. Gupta S., and Livne Z.A. (1988). Resolving a Conflict Situation with a Reference Outcome: An Axiomatic Model. *Management Science*, 34(11), s. 1303-1314.

12. Hakel M. (1968). How Often is Often? *American Psychologist*, 23, s. 533-534.
13. Hamm R.M. (1991). Selection of Verbal Probabilities: A Solution for Some Problems of Verbal Probability Expression. *Organizational Behavior and Human Decision Processes*, 48, s. 193-223.
14. Hartigan J. (1975). *Clustering Algorithms*. Wiley.
15. Huizingh, Eelko K.R.E. & Vrolijk, Hans C.J. (1997). A Comparison of Verbal and Numerical Judgments in the Analytic Hierarchy Process. *Organizational Behavior and Human Decision Processes*, Elsevier, 70(3), s. 237-247.
16. Jadidi O., Hong T.S., Firouzi F., Yusuff R.M. (2008). An Optimal Grey Based Approach Based on TOPSIS Concept for Supplier Selection Problem. *International Journal of Management Science and Engineering Management*, 4, 2, s. 104-117.
17. Kahraman C. (2008). *Fuzzy Multi-Criteria Decision Making. Theory and Application with Recent Development*. Series Springer. Optimization and its Applications, 16.
18. Keeney R., Raiffa H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press, New York.
19. Kersten G.E., Noronha S.J. (1999). WWW-Based Negotiation Support: Design, Implementation And Use. *Decision Support Systems*, 25, s. 135-154.
20. Kong A., Barnett G.O., Mosteller F., & Youtz C. (1986). How medical Professionals Evaluate Expressions of Probability. *The New England Journal of Medicine*, 315, s. 740-744.
21. E.F. Lane and Verdini W.A. (1989). A consistency Test for AHP Decision Makers. *Decision Science* 20, s. 575-590.
22. Mapes R.E.A. (1979). Verbal and Numerical Estimates of Probability in Therapeutic Contexts. *Social Science Medicine*, 13, s. 277-282.
23. Moshkovich H., Mechitov A.O.D. (2005). *Verbal Decision Analysis*. Springer, New York, s. 609-633.
24. Mousseau V., Slowinski R. (1998). Inferring an ELECTRE TRI Model from Assignment Examples. *J. of Global Optimization*, 12, s. 157-174.
25. Murphy C.K. (1993). Limits of the Analytical Hierarchy Process from its Consistency Index. *European Journal of Operational Research*, 65, s. 138-139.
26. Nakao M.A., & Axelrod S. (1983). Numbers are Better than Words: Verbal Specifications of Frequency Have No Place in Medicine, *The American Journal of Medicine*, 74, s. 1061-1063.
27. Parzen E. (1962). On Estimation of a Probability Density Function and Mode. *Ann. Math. Stat.*, 33, s. 1065-1076.
28. Pepper S., & Prytulak L.S. (1974). Sometimes Frequently Means Seldom: Context Effects in the Interpretation of Quantitative Expressions. *Journal of Research in Personality*, 8, s. 95-101.

29. Raiffa H., Richardson J., Metcalfe D. (2002). *Negotiation Analysis*. The Belknap Press of Harvard University Press, Cambridge.
30. Roszkowska E., Wachowicz T. (2011). Studium metodologiczne, analiza aplikacyjności oraz komputerowe wspomaganie metody TOPSIS. XXX konferencja naukowa MZBO'11. Trzebieiszowice, 16-18 października.
31. Roy B. (1991). The Outranking Approach and the Foundations of Electre Methods. *Theory and Decision*, 31, 1, s. 49-73.
32. Saaty T.L., Alexander J.M. (1989). *Conflict Resolution: The Analytic Hierarchy Approach*. Praeger, New York.
33. Saaty T. (1980). *The Analytic Hierarchy Process*. McGraw Hill, New York.
34. Teigen K.H., Brun W. (1999). The Directionality of Verbal Probability Expressions: Effects on Decisions, Predictions and Probabilistic Reasoning. *Organizational Behaviour and Human Decision Processes*, 80, s. 155-190.
35. Timmermans D. (1994). The Roles of Experience and Domain of Experiment in Using Numerical and Verbal Probability Terms in Medical Decisions. *Medical Decision Making*, 14(2), s. 146-156.
36. Wachowicz T. (2008). Metoda Even Swaps w tworzeniu systemu oceny ofert negocjacyjnych. W: *Metody i zastosowania badań operacyjnych '07*. Red. D. Kopańska-Bródka. AE, Katowice, s. 297-308.
37. Zimmer A.C. (1983). Verbal vs. Numerical Processing of Subjective Probabilities. In: *Decision making under uncer*, Ed. R.W. Scholz. 30, 328-330.

USING THE LINGUISTIC SCALES FOR DEFINING UTILITIES IN PREFERENCE ELICITATION PROCESS

Summary

In this work we discuss the problem of defining preferences by means of verbal scales. We propose a novel method combining two simple 7-point verbal scales to keep both the intuitive interpretation of the linguistic terms and adequate precision of the obtained scale. The solution we propose is applied in the negotiation context as a supportive tool of the NegoManage electronic negotiation system. In our paper we discuss pros and cons of using both numeric and linguistic scales. Finally we discuss the way the novel linguistic scale is built and described how the numerical equivalents are determined for each linguistic term of the proposed scale.

Hanna Bury, Dariusz Wagner

Instytut Badań Systemowych PAN w Warszawie

EFEKTYWNE METODY WYZNACZANIA MEDIANY KEMENY'EGO

Wprowadzenie

Istnieje wiele sposobów upraszczania zadania wyznaczania mediany Kemeny'ego, które w przypadku ogólnym stanowi problem NP-trudny. W pracy przedstawiono jedną z takich metod.

Założymy, że dany jest zbiór obiektów $\mathcal{O} = \{O_1, \dots, O_n\}$ oraz zbiór ekspertów $\mathcal{H} = \{1, \dots, K\}$. Zadaniem ekspertów jest uporządkowanie zbioru obiektów zgodnie z przyjętym kryterium lub zbiorem kryteriów. Zakładamy również, że wszystkie obiekty ze zbioru $\mathcal{O}$ są porównywane ze sobą oraz, że w opiniach ekspertów mogą występować obiekty równoważne. W dotychczasowych rozważaniach posługiwano się przede wszystkim macierzami porównań parami oraz wyznaczaną na ich podstawie tzw. macierzą strat R . Celem pracy jest zwrócenie uwagi na fakt, że można uzyskać interesujące wyniki wprowadzając do rozważań macierz rozkładu głosów ekspertów L .

Zaproponowana metoda wykorzystuje twierdzenie Litvaka [3] podające warunek konieczny i dostateczny, który musi spełniać uporządkowanie, aby stanowiło medianę. Można pokazać – wykorzystując to twierdzenie oraz zależności między macierzami R i L – że wyznaczanie mediany sprowadza się do wyznaczenia relacji większości, jaką powinny spełniać obiekty tworzące szukane uporządkowanie. Następnie należy dokonać eliminacji ze zbioru wszystkich możliwych uporządkowań n obiektów tych uporządkowań, w których wymagane relacje większości między obiektami nie są spełnione. Jeżeli w wyniku tego postępowania otrzymano więcej niż jedno uporządkowanie, należy wybrać to, którego odległość (w sensie Kemeny'ego) od zbioru uporządkowań podanych przez ekspertów jest najmniejsza.

1. Definicje mediany Kemeny'ego. Podejście klasyczne [1]

Oceny pary obiektów podane przez ekspertów mogą mieć postać (dla uproszczenia zapisu pominięto indeks eksperta)

- $O_i \succ O_j$, co oznacza, że obiekt O_i jest – ze względu na przyjęte kryterium (zbiór kryteriów) – lepszy od obiektu O_j ,
 $O_i \approx O_j$, co oznacza, że obiekty O_i oraz O_j są – ze względu na przyjęte kryterium (zbiór kryteriów) – równoważne,
 $O_i \prec O_j$, co oznacza, że obiekt O_i jest – ze względu na przyjęte kryterium (zbiór kryteriów) – gorszy od obiektu O_j .

Stosowany jest również zapis w postaci uporządkowania

$$P = \{O_{i_1}, O_{i_2}, \dots, (O_{i_s}, O_{i_t}), \dots, O_{i_{n-1}}, O_{i_n}\} \quad (1)$$

który oznacza, że obiekt O_{i_1} jest lepszy od obiektu O_{i_2} ; obiekty równoważne (O_{i_s}, O_{i_t}) są ujęte w nawiasach.

Dla ocen par obiektów, które podaje ekspert o numerze k , tworzymy macierz porównań obiektów parami, której elementy są określone następująco

$$A^k = [a_{ij}^k], \text{ gdzie } a_{ij}^k = \begin{cases} 1 & \text{jeżeli } O_i \succ^k O_j \\ 0 & \text{jeżeli } O_i \approx^k O_j, \\ -1 & \text{jeżeli } O_i \prec^k O_j \end{cases} \quad a_{ii} = 0, i = 1, \dots, n. \quad (2)$$

W dalszych rozważaniach przyjmujemy, że eksperci podają swoje oceny w postaci uporządkowań^{*}.

Zakładamy, że mamy uporządkowanie P^k podane przez eksperta o numerze k oraz dane uporządkowanie P .

Definicja 1 [1]

Odległość między parą obiektów O_i, O_j w uporządkowaniu P oraz parą tych samych obiektów w uporządkowaniu P^k jest definiowana następująco

$$d_{ij}(P, P^k) = |a_{ij} - a_{ij}^k| \quad (3)$$

Definicja 2 [1]

Odległość między uporządkowaniami P oraz P^k jest definiowana następująco

$$d(P, P^k) = \sum_{i=1}^{n-1} \sum_{j>i}^n d_{ij}(P, P^k) = \sum_{i=1}^{n-1} \sum_{j>i}^n |a_{ij} - a_{ij}^k| \quad (4)$$

^{*} Mając dane uporządkowanie można na jego podstawie uzyskać – w sposób jednoznaczny – macierz porównań obiektów parami. Należy podkreślić, że znajomość tej macierzy nie zawsze umożliwia wyznaczenie odpowiadającego jej uporządkowania. Taka sytuacja ma miejsce, gdy podane oceny par obiektów nie są przechodnie.

Przyjmujemy, że w uporządkowaniu P obiekt O_i poprzedza O_j , czyli $a_{ij}=1$. Wprowadzamy oznaczenie

$$r_{ij}^k = |a_{ij}^k - a_{ij}| = |a_{ij}^k - 1| \quad k = 1, \dots, K \quad (5)$$

Współczynniki r_{ij}^k mogą przyjmować następujące wartości

$$r_{ij}^k = |a_{ij}^k - 1| = \begin{cases} 0 & \text{jeżeli } a_{ij}^k = 1 \\ 1 & \text{jeżeli } a_{ij}^k = 0 \\ 2 & \text{jeżeli } a_{ij}^k = -1 \end{cases} \quad (6)$$

Wzór (4) przybiera zatem postać

$$d(P, P^k) = \sum_{i=1}^{n-1} \sum_{j>i}^n |a_{ij} - a_{ij}^k| = \sum_{i=1}^{n-1} \sum_{j>i}^n |a_{ij}^k - 1| = \sum_{i=1}^{n-1} \sum_{j>i}^n r_{ij}^k \quad (7)$$

$$\text{Wprowadzamy oznaczenie } r_{ij} = \sum_{k=1}^K r_{ij}^k. \quad (8)$$

Definicja 3 [3]

Odległość $d(P, P^{(k)})$ uporządkowania P od zbioru uporządkowań $P^{(k)} = \{P^1, \dots, P^K\}$ podanych przez ekspertów wynosi

$$d(P, P^{(k)}) = \sum_{k=1}^K d(P, P^k) = \sum_{k=1}^K \sum_{i=1}^{n-1} \sum_{j>i}^n r_{ij}^k = \sum_{i=1}^{n-1} \sum_{j>i}^n r_{ij} \quad (9)$$

Macierz $R=[r_{ij}]$ nosi nazwę macierzy strat. Ze wzoru (6) wynika, że wartości elementów macierzy R zależą wyłącznie od postaci uporządkowań podanych przez ekspertów.

Definicja 4 [3]

Medianą Kemeny'ego nazywamy takie uporządkowanie $\tilde{P}$, że

$$d(\tilde{P}, P^{(k)}) = \min_P d(P, P^{(k)}) \quad (10)$$

Z definicji mediany Kemeny'ego (10) oraz wzoru (9) wynika, że zadanie wyznaczenia mediany Kemeny'ego można sprowadzić do takiego przedstawienia wierszy i kolumn macierzy R (tzn. zmiany kolejności obiektów), aby suma elementów r_{ij} nad przekątną główną była jak najmniejsza. Należy podkreślić, że istnieje wiele sposobów upraszczania zadania wyznaczenia mediany Kemeny'ego. Niektóre z nich zostaną omówione w niniejszej pracy.

Definicja 5 [3]

Macierz strat jest przechodnia, jeżeli z warunku

$$r_{i_j i_{j+1}} \leq r_{i_{j+1} i_j} \text{ oraz } r_{i_{j+1} i_{j+s}} \leq r_{i_{j+s} i_{j+1}} \text{ wynika, że } r_{i_j i_{j+s}} \leq r_{i_{j+s} i_j} \quad (11)$$

$$i = 1, \dots, n, j = 1, \dots, n-1, s = 2, \dots, (n-j)$$

Warto odnotować, że warunek przechodniości nie wymaga aby $s=2$. W przypadku ogólnym dla $K \geq 3$ macierz strat odpowiadająca dowolnym uporządkowaniom $P^{(k)} = \{P^1, \dots, P^K\}$ nie musi być przechodnia.

Definicja 6 [3]

Przyjmujemy, że zwycięzcą w sensie Condorceta jest obiekt, który zdaniem większości ekspertów ($K_W \geq K/2$) poprzedza pozostałe obiekty^{*}.

Definicja 7 [3]

Zbiór uporządkowań ma własność Condorceta, jeżeli dla każdego podzbioru obiektów istnieje zwycięzca w sensie Condorceta.

Obiekty będące zwycięzcami w sensie Condorceta są wyznaczane następująco. Najpierw określa się zwycięzcę Condorceta dla całego zbioru obiektów $\{O_i\}$ $i = 1, \dots, n$. Następnie obiekt ten jest usuwany ze zbioru obiektów i wyznaczany jest – o ile istnieje – zwycięzca Condorceta dla $(n-1)$ elementowego podzbioru obiektów. Tak wyznaczony obiekt jest usuwany ze zbioru obiektów i postępowanie jest powtarzane dla kolejnych podzbiorów.

Twierdzenie 1 [3]

Macierz strat odpowiadająca uporządkowaniom $P^{(k)} = \{P^1, \dots, P^K\}$ jest przechodnia wtedy i tylko wtedy, gdy zbiór tych uporządkowań ma własność Condorceta.

Twierdzenie 2 [3]

Jeżeli dany zbiór uporządkowań $P^{(k)} = \{P^1, \dots, P^K\}$ ma własność Condorceta, to medianę Kemeny'ego stanowi uporządkowanie obiektów $\tilde{P} = \{O_{i_1}, O_{i_2}, \dots, O_{i_{n-1}}, O_{i_n}\}$ będących zwycięzcami Condorceta dla kolejnych podzbiorów obiektów. Dla tego uporządkowania odległość $d(\tilde{P}, P^{(k)})$ jest kresem dolnym odległości od zbioru uporządkowań $P^{(k)}$. Oznacza się ją literą H .

^{*} Niektórzy autorzy przyjmują, że warunek ten ma postać ($K_W > K/2$).

Twierdzenie 3 [3]

Jeżeli w medianie nie występują obiekty równoważne, wówczas

$$d(P, P^{(k)}) = \sum_{(i,j) \in I_P^{(1)}} r_{ij} \quad (12)$$

gdzie $I_P^{(1)}$ – zbiór indeksów (i,j) , dla których w uporządkowaniu P mamy $O_i \succ O_j$, tzn. $a_{ij} = 1$.

Ta zależność może być zastosowana do wyznaczenia kresu dolnego H odległości (9). Litvak wykazał, że spełnione jest następujące twierdzenie.

Twierdzenie 4 [3]

$$H = \sum_{i=1}^{n-1} \sum_{j>i}^n \min(r_{ij}, r_{ji}) \quad (13)$$

2. Zastosowanie macierzy rozkładu głosów ekspertów

W dotychczasowych rozważaniach posługiwano się przede wszystkim macierzą porównań parami. Wprowadzając do rozważań macierz rozkładu głosów ekspertów można uzyskać istotne wyniki.

Założymy, że dany jest zbiór uporządkowań n obiektów podanych przez K ekspertów $P^{(k)} = \{P^1, \dots, P^K\}$. Założymy również, że w opiniach ekspertów mogą występować obiekty równoważne.

Wprowadzimy oznaczenia:

$$\begin{aligned} l_{ij} & - \text{liczba ekspertów, których zdaniem } O_i \succ O_j \\ l_{ji} & - \text{liczba ekspertów, których zdaniem } O_j \succ O_i \\ m_{ij} & - \text{liczba ekspertów, których zdaniem } O_i \approx O_j, \text{ tzn., że obiekty } O_i \text{ oraz } O_j \text{ są równoważne w sensie przyjętego kryterium (zbioru kryteriów)}. \end{aligned} \quad (14)$$

$$\text{Zatem spełniona jest równość } l_{ij} + l_{ji} + m_{ij} = K. \quad (15)$$

Z definicji macierzy strat $R = [r_{ij}]$ (6), (8), wynika, że

$$r_{ij} = \sum_k |a_{ij}^k - 1|_{O_i \succ^k O_j} + \sum_k |a_{ij}^k - 1|_{O_i \approx^k O_j} + \sum_k |a_{ij}^k - 1|_{O_i \prec^k O_j} \quad (16)$$

skąd

$$r_{ij} = 0 \cdot l_{ij} + 1 \cdot m_{ij} + 2l_{ji} \quad (17)$$

Jeżeli w uporządkowaniach występują obiekty równoważne zazwyczaj przyjmuje się – aby był spełniony wzór (15) – że każdy z obiektów równoważnych otrzymuje 1/2 głosu eksperta.

Niech

$$\tilde{l}_{ij} = l_{ij} + 0.5m_{ij} \quad (18)$$

Wówczas współczynniki macierzy strat można zapisać w postaci

$$r_{ij} = 2(l_{ji} + 0.5 \cdot m_{ij}) = 2\tilde{l}_{ji} \quad (19)$$

Symbole $\tilde{l}_{ij}$ oraz l_{ij} będą używane zamiennie, ponieważ sposób wyznaczania współczynników macierzy rozkładu głosów ekspertów zazwyczaj wynika z kontekstu. Mamy więc zależność wiążącą macierze strat i rozkładu głosów ekspertów

$$R = 2L^T \quad (20)$$

Zależność ta umożliwia sformułowanie podanych przez Litvaka [3] twierdzeń dotyczących własności macierzy strat oraz wyznaczania kresu dolnego odległości za pomocą macierzy rozkładu głosów ekspertów.

$$\text{Wzór (13) może być zapisany w postaci } H = 2 \sum_{i=1}^{n-1} \sum_{j>i}^n \min(l_{ij}, l_{ji}) \quad (21)$$

Odejmując od wyrażenia (17) równość (15) otrzymujemy

$$r_{ij} - K = 2l_{ji} + 1 \cdot m_{ij} - l_{ij} - l_{ji} - m_{ij} = l_{ji} - l_{ij} = -\Delta l_{ij} \quad (22)$$

$$\text{Tak więc } r_{ij} = K - \Delta l_{ij} \quad (23)$$

$$\text{Ze wzoru (23) wynika także, że } |r_{ij} - r_{ji}| = 2|\Delta l_{ij}| \quad (24)$$

Jeżeli $\Delta l_{ij} \neq 0$, to $|r_{ij} - r_{ji}|_{\min} = 2$, bowiem $|\Delta l_{ij}|_{\min} = 1$. Zatem minimalna zmiana odległości między uporządkowaniami wynosi 2.

3. Wyznaczanie mediany Kemeny'ego

W pracy Litvaka [3] jest podane twierdzenie mające duże znaczenie praktyczne. W dalszych rozważaniach zostanie pokazane, jak zastosowanie tego twierdzenia pozwala istotnie uprościć proces wyznaczania mediany Ke-

meny'ego. Różne wersje tego twierdzenia były podawane w kilku pracach, np. Saari, Merlin [4], Truchon [5]. Jednakże ich autorzy nie odwoływali się do oryginalnej pracy Litvaka opublikowanej znacznie wcześniej, bo w 1982 roku.

Twierdzenie 5 [3]

Założymy, że jest dany zbiór uporządkowań $P^{(k)} = \{P^1, \dots, P^K\}$, dla którego wyznaczono macierz strat $R = [r_{ij}]$, $i, j = 1, \dots, n$. Jeżeli uporządkowanie $\tilde{P}$ stanowi medianę tego zbioru uporządkowań, to jest spełniony warunek

$$r_{i_v i_{v+1}} \leq r_{i_{v+1} i_v}, \quad v = 1, \dots, n-1 \quad (25)$$

Jeżeli dla danego uporządkowania P jest spełniony warunek (25) a macierz strat R jest przechodnia, to uporządkowanie P stanowi medianę zbioru uporządkowań $P^{(k)}$.

Aby zapisać warunek (25) używając elementów macierzy rozkładu głosów ekspertów, przywołamy wzór (23)

$$r_{i_v i_{v+1}} = K - (l_{i_v i_{v+1}} - l_{i_{v+1} i_v}) \quad \text{oraz} \quad r_{i_{v+1} i_v} = K - (l_{i_{v+1} i_v} - l_{i_v i_{v+1}}) \quad (26)$$

Warunek (25) można zatem zapisać jak następuje

$$K - (l_{i_v i_{v+1}} - l_{i_{v+1} i_v}) \leq K - (l_{i_{v+1} i_v} - l_{i_v i_{v+1}}) \quad (27)$$

$$\text{skąd} \quad 0 \leq 2(l_{i_v i_{v+1}} - l_{i_{v+1} i_v}) \quad (28)$$

$$\text{i ostatecznie} \quad 0 \leq (l_{i_v i_{v+1}} - l_{i_{v+1} i_v}) \quad (29)$$

Warunek (25) jest warunkiem koniecznym, aby uporządkowanie $\tilde{P}$ było medianą. Zatem (29) jest również warunkiem koniecznym, który musi spełniać mediana. Konsekwencją tego warunku jest wymaganie, aby relacje większości między parami obiektów – wynikające z macierzy rozkładu głosów ekspertów – były zachowane w medianie. Zatem na podstawie wzoru (29) można sformułować następujący wniosek: uporządkowania, w których nie są spełnione relacje większości wynikające z macierzy rozkładu głosów ekspertów, nie mogą być medianą.

Podejście, którego podstawę stanowi to stwierdzenie, zastosował M. Truchon [5] do uproszczenia procesu wyznaczania mediany Kemeny'ego.

Kolejność postępowania jest następująca:

1°. Na podstawie macierzy rozkładu głosów ekspertów wyznacza się pary obiektów spełniające relację większości.

2°. Dla par obiektów wyznaczonych w punkcie 1° określa się relację przeciwną, tzn., jeżeli między parą obiektów O_i oraz O_j zachodzi relacja większości $O_i \succ O_j$, to przyjmujemy, że $O_j \succ O_i$.

3°. Ustalamy wszystkie możliwe uporządkowania, które mogą stanowić medianę. W przypadku n obiektów i przy założeniu, że w medianie nie występują obiekty równoważne, należy rozpatrzyć $n!$ uporządkowań.

4°. Ze zbioru uporządkowań wyznaczonych w punkcie 3° usuwamy te, w których między sąsiednimi obiektami występują relacje ustalone w punkcie 2°.

5°. Dla uporządkowań, które nie zostały usunięte w kroku 4° sprawdzamy, dla którego uporządkowania (dla których uporządkowań) odległość (9) jest najmniejsza. Uporządkowania, które spełniają ten warunek stanowią medianę Kemeny'ego.

Sprawdzenia, czy jest spełniony warunek z pkt. 5° można dokonać na kilka sposobów, w szczególności mając wyznaczoną macierz rozkładu głosów ekspertów L można dla każdego uporządkowania P obliczyć odległość $d(P, P^{(k)})$ zgodnie z wzorem

$$d(P, P^{(k)}) = \sum_{i=1}^{n-1} \sum_{j>i}^n r_{ij} = \sum_{i=1}^{n-1} \sum_{j>i}^n [K - (l_{ij} - l_{ji})] = \frac{n(n-1)}{2} K - \sum_{i=1}^{n-1} \sum_{j>i}^n \Delta l_{ij} \quad (30)$$

W przedstawionej procedurze – służącej do uproszczenia procesu wyznaczania mediany Kemeny'ego – zakłada się eliminację tych uporządkowań, w których sąsiadujące obiekty nie spełniają zasady większości wynikającej z macierzy rozkładu głosów ekspertów.

Jednakże może zdarzyć się sytuacja, że w uporządkowaniach pozostałych po tej eliminacji mogą występować pary obiektów niezajmujących sąsiednich pozycji, które również nie spełniają relacji większości wynikających z macierzy rozkładu głosów ekspertów. Dla tych par obiektów $\Delta l_{ij} < 0$.

Zatem zgodnie z wzorem (30), odległość $d(P, P^{(k)})$ będzie tym mniejsza im suma $\sum_{i=1}^{n-1} \sum_{j>i}^n \Delta l_{ij}$ będzie mieć większą wartość. Ze względu na to, że dla rozpatrywanych par obiektów $\Delta l_{ij} < 0$, odpowiadające im składniki sumy będą powiększać wartość odległości. Można zatem przyjąć, że uporządkowanie, w którym suma $|\Delta l_{ij}|$ wyznaczona dla par obiektów nie spełniających relacji większości wynikających z macierzy rozkładu głosów ekspertów jest naj-

mniejsza, stanowi medianę Kemeny'ego. Tak właśnie jest formułowana zasada wyznaczania mediany Kemeny'ego w pracach Klamlera [2] oraz Saariego i Merlina [4].

4. Przykłady

Aby zilustrować przedstawiony w pracy algorytm wyznaczania mediany Kemeny'ego, którego podstawę stanowi twierdzenie 5, podamy jego zastosowanie do konkretnego przykładu.

Przykład 1

Dany jest zestaw pięciu uporządkowań zbioru czterech obiektów.

$$\begin{aligned}
 P^1: & O_1, O_2, O_4, O_3 \\
 P^2: & O_4, O_2, O_1, O_3 \\
 P^3: & O_2, O_3, O_4, O_1 \\
 P^4: & O_3, O_4, O_1, O_2 \\
 P^5: & O_1, O_2, O_3, O_4
 \end{aligned} \tag{31}$$

Macierz rozkładu głosów ekspertów

$$L = \begin{array}{c|cccc} & O_1 & O_2 & O_3 & O_4 \\ \hline O_1 & - & 3 & 3 & 2 \\ \hline O_2 & 2 & - & 4 & 3 \\ \hline O_3 & 2 & 1 & - & 3 \\ \hline O_4 & 3 & 2 & 2 & - \end{array} \tag{32}$$

Relacja większości jest spełniona jeżeli $l_{ij} \geq 2\frac{1}{2}$.

Pary obiektów, które spełniają relację większości są następujące

$$O_1 \succ O_2; O_1 \succ O_3; O_2 \succ O_3; O_2 \succ O_4; O_3 \succ O_4; O_4 \succ O_1 \tag{33}$$

Przyjmujemy, że mediany dla zbioru uporządkowań będziemy szukać wśród uporządkowań nie zawierających obiektów równoważnych.

Wiadomo, że dla 4 obiektów takich uporządkowań jest $4!=24$. Są one, jak następuje (dla uproszczenia zapisu pominięto symbol $\succ$).

1° $O_1 O_2 O_3 O_4$	9° $O_2 O_3 O_1 O_4$	17° $O_3 O_4 O_1 O_2$
2° $O_1 O_2 O_4 O_3$	10° $O_2 O_3 O_4 O_1$	18° $O_3 O_4 O_2 O_1$
3° $O_1 O_3 O_2 O_4$	11° $O_2 O_4 O_1 O_3$	19° $O_4 O_1 O_2 O_3$
4° $O_1 O_3 O_4 O_2$	12° $O_2 O_4 O_3 O_1$	20° $O_4 O_1 O_3 O_2$
5° $O_1 O_4 O_2 O_3$	13° $O_3 O_1 O_2 O_4$	21° $O_4 O_2 O_1 O_3$
6° $O_1 O_4 O_3 O_2$	14° $O_3 O_1 O_4 O_2$	22° $O_4 O_2 O_3 O_1$
7° $O_2 O_1 O_3 O_4$	15° $O_3 O_2 O_1 O_4$	23° $O_4 O_3 O_1 O_2$
8° $O_2 O_1 O_4 O_3$	16° $O_3 O_2 O_4 O_1$	24° $O_4 O_3 O_2 O_1$

(34)

Medianą nie mogą być uporządkowania, w których relacje między sąsiednimi obiektami są przeciwne do podanych w (33). Należy zatem wykluczyć uporządkowania, w których

$$\begin{array}{lll}
 \text{a) } O_1 \succ O_4; & \text{b) } O_2 \succ O_1; & \text{c) } O_3 \succ O_1 \\
 \text{d) } O_3 \succ O_2; & \text{e) } O_4 \succ O_2; & \text{f) } O_4 \succ O_3
 \end{array} \quad (35)$$

Z analizy 24 uporządkowań wynika, że:

- warunek a) wyklucza uporządkowania 5°, 6°, 8°, 9°, 14°, 15°
- warunek b) wyklucza uporządkowania 7°, 8°, 15°, 18°, 21°, 24°
- warunek c) wyklucza uporządkowania 9°, 12°, 13°, 14°, 22°, 23°
- warunek d) wyklucza uporządkowania 3°, 6°, 15°, 16°, 20°, 24°
- warunek e) wyklucza uporządkowania 4°, 5°, 14°, 18°, 21°, 22°
- warunek f) wyklucza uporządkowania 2°, 6°, 8°, 12°, 23°, 24°

(36)

Warunek (33) spełniają jedynie uporządkowania

$$\begin{array}{lll}
 1^\circ & O_1, O_2, O_3, O_4 & 10^\circ & O_2, O_3, O_4, O_1 & 11^\circ & O_2, O_4, O_1, O_3 \\
 17^\circ & O_3, O_4, O_1, O_2 & 19^\circ & O_4, O_1, O_2, O_3
 \end{array} \quad (37)$$

Aby wyznaczyć medianę, wystarczy sprawdzić, dla którego z tych uporządkowań odległość $d(P, P^{(k)})$ osiąga wartość minimalną lub wyrażenie

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n \Delta_{ij} \text{ osiąga wartość maksymalną.}$$

Macierze $[\Delta l_{ij}]$ dla rozpatrywanych uporządkowań mają postać

	O_1	O_2	O_3	O_4	$\sum_{j>i} \Delta l_{ij}$
O_1	–	1	1	–1	1
O_2		–	3	1	4
O_3			–	1	1
O_4				–	6

	O_2	O_3	O_4	O_1	$\sum_{j>i} \Delta l_{ij}$
O_2	–	3	1	–1	3
O_3		–	1	–1	0
O_4			–	1	1
O_1				–	4

	O_2	O_4	O_1	O_3	$\sum_{j>i} \Delta l_{ij}$
O_2	–	1	–1	3	3
O_4		–	1	–1	0
O_1			–	1	1
O_3				–	4

	O_3	O_4	O_1	O_2	$\sum_{j>i} \Delta l_{ij}$
O_3	–	1	–1	–3	–3
O_4		–	1	–1	0
O_1			–	1	1
O_2				–	–2

	O_4	O_1	O_2	O_3	$\sum_{j>i} \Delta l_{ij}$
O_4	–	1	–1	–1	–1
O_1		–	1	1	2
O_2			–	3	3
O_3				–	4

(38)

Z porównania macierzy $[\Delta l_{ij}]$ wynika, że medianę stanowi uporządkowanie O_1, O_2, O_3, O_4 .

Jego odległość od zbioru uporządkowań wyznaczona za pomocą wzoru (30) wynosi $\frac{4(4-1)}{2}5 - \sum_i \sum_j \Delta l_{ij} = 30 - 6 = 24$. Kres dolny macierzy

strat określonej dla uporządkowań (1) $H = 22$. Nie istnieje uporządkowanie dla którego $d(P, P^{(k)}) = H$. A zatem uporządkowanie O_1, O_2, O_3, O_4 jest medianą, ponieważ odpowiadająca mu odległość od kresu dolnego jest najmniejsza (i równa 2).

Przykład 2

Założmy, że jest dana macierz rozkładu głosów ekspertów wyznaczona na podstawie uporządkowań 5 obiektów, podanych przez 7 ekspertów.

$$L = \begin{array}{c|ccccc} & O_1 & O_2 & O_3 & O_4 & O_5 \\ \hline O_1 & 0 & 2.5 & 3 & 4 & 5 \\ \hline O_2 & 4.5 & 0 & 3 & 5 & 5 \\ \hline O_3 & 4 & 4 & 0 & 4.5 & 5 \\ \hline O_4 & 3 & 2 & 2.5 & 0 & 4.5 \\ \hline O_5 & 2 & 2 & 2 & 2.5 & 0 \end{array} \quad K_w \geq \frac{7}{2} = 3.5 \quad (39)$$

Z macierzy L wynika, że obiekt O_3 jest zwycięzcą w sensie Condorceta.

Po wykreśleniu go z macierzy (39) otrzymujemy

$$(40) \quad \begin{array}{c|cccc} & O_1 & O_2 & O_4 & O_5 \\ \hline O_1 & 0 & 2.5 & 4 & 5 \\ \hline O_2 & 4.5 & 0 & 5 & 5 \\ \hline O_4 & 3 & 2 & 0 & 4.5 \\ \hline O_5 & 2 & 2 & 2.5 & 0 \end{array}$$

$$(41) \quad \begin{array}{c|ccc} & O_1 & O_4 & O_5 \\ \hline O_1 & 0 & 4 & 5 \\ \hline O_4 & 3 & 0 & 4.5 \\ \hline O_5 & 2 & 2.5 & 0 \\ \hline & & & \end{array}$$

Kolejnym zwycięzcą w sensie Condorceta jest obiekt O_2 .

Po wykreśleniu go z macierzy (40) otrzymujemy macierz (41).

Następnymi zwycięzcami w sensie Condorceta są obiekty O_1 oraz O_4 .

Uporządkowanie wynikowe ma postać O_3, O_2, O_1, O_4, O_5 . (42)

Macierz $[\Delta l_{ij}]$ uporządkowana w podanej kolejności obiektów jest następująca

$$(43) \quad \begin{array}{c|ccccc|c} & O_3 & O_2 & O_1 & O_4 & O_5 & \sum_{j>i} \Delta l_{ij} \\ \hline O_3 & - & 1 & 1 & 2 & 3 & 7 \\ \hline O_2 & & - & 2 & 3 & 3 & 8 \\ \hline O_1 & & & - & 1 & 3 & 4 \\ \hline O_4 & & & & - & 2 & 2 \\ \hline O_5 & & & & & - & 21 \end{array}$$

Wykorzystując wzór (30) otrzymujemy, że dla uporządkowania (42) odległość $d(P, P^{(k)}) = 7 \frac{5(5-1)}{2} - 21 = 49$.

Kres dolny odległości $d(P, P^{(k)})$ obliczony na podstawie wzoru (21) wynosi 49. A zatem uporządkowanie (42) jest medianą.

Warto podkreślić, że w niektórych przypadkach proces wyznaczania mediany Kemeny'ego można istotnie uprościć, nawet wtedy, gdy macierz strat nie jest przechodnia. Sposób postępowania przedstawiono na poniższym przykładzie.

Przykład 3

Załóżmy, że jest dana macierz rozkładu głosów ekspertów wyznaczona na podstawie uporządkowań 6 obiektów, podanych przez 11 ekspertów.

	O_1	O_2	O_3	O_4	O_5	O_6
O_1	–	7	7	6	6	7
O_2	4	–	7	4	9	6
O_3	4	4	–	4	5	5
O_4	5	7	7	–	7	5
O_5	5	2	6	4	–	5
O_6	4	5	6	6	6	–

(44)

Liczba ekspertów wynosi 11, czyli większość $K_w \geq 11/2 = 5.5$. Pary obiektów (O_i, O_j) spełniające relację większości zacięniowano. Z macierzy (44) wynika, że obiekt O_1 jest zwycięzcą w sensie Condorceta. Po wykreśleniu go z macierzy (44) otrzymujemy

	O_2	O_3	O_4	O_5	O_6
O_2	–	7	4	9	6
O_3	4	–	4	5	5
O_4	7	7	–	7	5
O_5	2	6	4	–	5
O_6	5	6	6	6	–

(45)

W macierzy tej nie występuje zwycięzca w sensie Condorceta, ale przyjmujemy, że obiekt który jest bliski zwycięzcy Condorceta zostaje uznany za tego zwycięzcę. Z macierzy (45) wynika, że może to być obiekt O_2 , O_4 lub O_6 . Po wykreśleniu obiektu O_4 otrzymujemy

	O_2	O_3	O_5	O_6
O_2	–	7	9	6
O_3	4	–	5	5
O_5	2	6	–	5
O_6	5	6	6	–

(46)

W macierzy (46) zwycięzcą w sensie Condorceta jest obiekt O_2 . Po wykreśleniu go z (46) otrzymujemy

	O_3	O_5	O_6
O_3	–	5	5
O_5	6	–	5
O_6	6	6	–

(47)

W macierzy (47) zwycięzcą w sensie Condorceta jest obiekt O_6 .

Po wykreśleniu go z (47) otrzymujemy

	O_3	O_5
O_3	–	5
O_5	6	–

(48)

W macierzy (48) zwycięzcą w sensie Condorceta jest obiekt O_5 . Uporządkowanie utworzone z kolejnych zwycięzców w sensie Condorceta ma więc postać

$$O_1, O_4, O_2, O_6, O_5, O_3 \quad (49)$$

Macierz $[\Delta l_{ij}]$ uporządkowana w podanej kolejności obiektów jest następująca

	O_1	O_4	O_2	O_6	O_5	O_3	$\sum_{j>i} \Delta l_{ij}$
O_1	–	1	3	3	1	3	11
O_4	–1	–	3	–1	3	3	8
O_2	–3	–3	–	1	7	3	11
O_6	–3	1	–1	–	1	1	2
O_5	–1	–3	–7	–1	–	1	1
O_3	–3	–3	–3	–1	–1	–	33

(50)

Wykorzystując wzór (30) otrzymujemy, że dla uporządkowania (49) odległość $d(P, P^{(k)}) = 11 \frac{6(6-1)}{2} - 33 = 132$.

Kres dolny odległości $d(P, P^{(k)})$ obliczony na podstawie wzoru (21) wynosi 130. Nie istnieje uporządkowanie dla którego $d(P, P^{(k)}) = H$. Zatem uporządkowanie (49) jest medianą, ponieważ odpowiadająca mu odległość od kresu dolnego jest najmniejsza (i równa 2).

Usuując z macierzy (45) obiekt O_6 otrzymujemy

	O_2	O_3	O_4	O_5
O_2	–	7	4	9
O_3	4	–	4	5
O_4	7	7	–	7
O_5	2	6	4	–

(51)

W macierzy (51) zwycięzcą w sensie Condorceta jest obiekt O_4 . Po wykreśleniu go z (51) otrzymujemy

	O_2	O_3	O_5
O_2	–	7	9
O_3	4	–	5
O_5	2	6	–

(52)

Kolejnymi zwycięzcami w sensie Condorceta są obiekty O_2 i O_5 . Uporządkowanie utworzone z kolejnych zwycięzców w sensie Condorceta ma więc postać

$$O_1, O_6, O_4, O_2, O_5, O_3 \quad (53)$$

Dla uporządkowania (53) odległość $d(P, P^{(k)})$ wyznaczona za pomocą wzoru (30) wynosi 132 a więc uporządkowanie to jest również medianą.

Usuując z macierzy (52) obiekt O_2 i następnie obiekty O_6 , O_4 i O_5 będące kolejnymi zwycięzcami w sensie Condorceta otrzymujemy uporządkowanie

$$O_1, O_2, O_6, O_4, O_5, O_3 \quad (54)$$

Dla uporządkowania (54) odległość $d(P, P^{(k)})$ wyznaczona za pomocą wzoru (30) wynosi 136, a więc uporządkowanie to nie jest medianą.

Przykład 4

Mamy dane uporządkowania 7 obiektów podane przez 11 ekspertów. Są one, jak następuje:

$$\begin{aligned}
P^1 &: (O_2, O_6), (O_5, O_7), O_3, O_4, O_1 \\
P^2 &: O_6, (O_1, O_3, O_4), O_5, O_7, O_2 \\
P^3 &: (O_5, O_6), O_2, (O_1, O_4), (O_3, O_7) \\
P^4 &: O_5, (O_1, O_3), (O_2, O_4), O_6, O_7 \\
P^5 &: O_7, O_4, (O_3, O_6), (O_1, O_2), O_5 \\
P^6 &: O_4, (O_5, O_6, O_7), (O_1, O_2), O_3 \\
P^7 &: O_4, O_5, (O_3, O_7), (O_1, O_2, O_6) \\
P^8 &: O_2, O_5, O_7, (O_4, O_6), (O_1, O_3) \\
P^9 &: O_6, (O_2, O_5, O_7), (O_1, O_3, O_4) \\
P^{10} &: O_1, O_2, O_4, (O_3, O_6), O_5, O_7 \\
P^{11} &: (O_1, O_2), O_3, (O_4, O_5), (O_6, O_7)
\end{aligned} \tag{55}$$

Liczba głosów wymagana przez zasadę większości $K_W \geq \frac{11}{2} = 5.5$.

Macierz rozkładu głosów z uwzględnieniem połówek głosów dla obiektów równoważnych jest następująca

	O_1	O_2	O_3	O_4	O_5	O_6	O_7
O_1	0	5	6	4.5	4	3.5	5
O_2	6	0	7	7	5.5	5	6.5
O_3	5	4	0	4	4	4	5
O_4	6.5	4	7	0	5.5	6	7
O_5	7	5.5	7	5.5	0	5	8.5
O_6	7.5	6	7	5	6	0	7
O_7	6	4.5	6	4	2.5	4	0

(56)

Pary obiektów, dla których zachodzi relacja większości zacięniowano.

Przyjmujemy, że mediany dla zbioru uporządkowań będziemy szukać wśród uporządkowań nie zawierających obiektów równoważnych.

Wiadomo, że dla 7 obiektów takich uporządkowań jest $7! = 5040$. Medianą nie mogą być uporządkowania, w których relacje między sąsiednimi obiektami są przeciwne do zacięniowanych. Stosując procedurę wykluczania uporządkowań opisaną w przykładzie 1 otrzymano 9 uporządkowań.

$$\begin{array}{cccc}
O_2, & O_5, & O_4, & O_6, & O_7, & O_1, & O_3 \\
O_2, & O_4, & O_6, & O_5, & O_7, & O_1, & O_3 \\
O_4, & O_6, & O_2, & O_5, & O_7, & O_1, & O_3 \\
O_4, & O_6, & O_5, & O_2, & O_7, & O_1, & O_3 \\
O_5, & O_2, & O_4, & O_6, & O_7, & O_1, & O_3 \\
O_5, & O_4, & O_6, & O_2, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_4, & O_5, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_5, & O_4, & O_7, & O_1, & O_3 \\
O_6, & O_5, & O_2, & O_4, & O_7, & O_1, & O_3
\end{array} \quad (57)$$

Dla czterech z nich

$$\begin{array}{cccc}
O_2, & O_4, & O_6, & O_5, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_5, & O_4, & O_7, & O_1, & O_3 \\
O_6, & O_5, & O_2, & O_4, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_4, & O_5, & O_7, & O_1, & O_3
\end{array} \quad (58)$$

$$\text{odległość } d(P, P^{(k)}) = 11 \frac{7(7-1)}{2} - 43 = 188.$$

Kres dolny odległości $d(P, P^{(k)})$ obliczony na podstawie wzoru (30) wynosi 186 i nie istnieje uporządkowanie dla którego $d(P, P^{(k)}) = H$. Zatem uporządkowania (58) są medianą, ponieważ odpowiadająca im odległość od kresu dolnego jest najmniejsza (i równa 2).

Podsumowanie

Zależność (20) wiążąca macierz strat R z macierzą rozkładu głosów ekspertów L umożliwia bezpośrednie zastosowanie tej ostatniej do wyznaczania mediany Kemeny'ego. Warunek konieczny (29), aby – wynikające z macierzy L – relacje większości między obiektami były zachowane w medianie, umożliwia eliminację ze zbioru rozważanych $n!$ uporządkowań tych uporządkowań, w których te relacje nie są spełnione. Podejście to w istotny sposób ogranicza licznosc zbioru uporządkowań przeszukiwanych przy stosowaniu przeglądu zupełnego i może być wykorzystane do efektywnego wyznaczania mediany Kemeny'ego.

Ponadto, w niektórych przypadkach, nawet gdy zbiór danych uporządkowań nie ma własności Condorceta, proces wyznaczania mediany Kemeny'ego można znacząco uprościć przyjmując obiekt bliski zwycięzcy Condorceta za tego zwycięzcę w celu utworzenia ciągu kolejnych zwycięzców

Condorceta. Zastosowanie macierzy rozkładu głosów ekspertów L jest istotne również z uwagi na fakt, że niekiedy macierz ta stanowi jedyne źródło informacji o preferencjach ekspertów (macierze porównań parami mogą nie być dostępne).

Literatura

1. Kemeny J.G., Snell L.J. (1962). Preference Ranking: An Axiomatic Approach. In: J.G. Kemeny and L.J. Snell: Mathematical Models in the Social Sciences. Ginn, New York.
2. Klamler C. (2004). The Dodgson Ranking and its Relation to Kemeny's Method and Slater's Rule. *Social Choice and Welfare* 23, s. 91-102.
3. Litvak B.G. (1982). Ekspertnaja informacija. Metody połączienija i analiza. Radio i Swjaz, Moskwa.
4. Saari D.G., Merlin V.R. (2000). A Geometric Examination of Kemeny's Rule. *Social Choice and Welfare*, 17, s. 403-438.
5. Truchon M. (2004). Aggregation of Rankings in Figure Skating. Cahier de Recherche 04-14. CIPANO. CIRPEE and Departement d'economique, Universite Laval, Quebec, Canada.

EFFECTIVE SOLUTION OF KEMENY'S MEDIAN

Summary

In the paper a heuristic procedure of determining the Kemeny median is presented. It is assumed that all the alternatives are compared and there are no ties in group judgement. The outranking matrix is applied to derive the solution. The method presented is based on the Litvak necessary and sufficient condition for an order to be the median. It is assumed that orders where adjacent alternatives do not fulfill majority rule should be eliminated. If the procedure ends with more than one order, the closest one (in the Kemeny sense) to the set of experts' orders should be chosen. This approach significantly reduces the number of preference orders to be searched for and may be efficiently applied to determine the Kemeny median.

Renata Dudzińska-Baryła

Uniwersytet Ekonomiczny w Katowicach

WYBRANE METODY SZACOWANIA FUNKCJI UŻYTECZNOŚCI I FUNKCJI WARTOŚCI

Wprowadzenie

W wielu pracach naukowych przyjmuje się postaci zarówno funkcji użyteczności, jak i funkcji wartości, zaproponowane przez twórców tych teorii wraz z parametrami przez nich oszacowanymi. W niektórych pracach proponowane są metody szacowania parametrów tych funkcji na podstawie eksperymentów przeprowadzanych wśród wybranej grupy osób. Postaci i parametry tych funkcji wpływają na wartość oceny wariantów decyzyjnych. Zatem właściwe ich oszacowanie ma istotne znaczenie w opisie procesu oceny i wyboru spośród niepewnych wariantów decyzyjnych w rzeczywistych sytuacjach decyzyjnych.

Celem pracy jest przedstawienie i porównanie metod szacowania funkcji użyteczności oraz funkcji wartości (wykorzystywanej w teorii perspektywy). Zaprezentowana zostanie istota klasycznych metod ekwiwalentu pewności i ekwiwalentu prawdopodobieństwa oraz metoda gamble-tradeoff służąca do szacowania funkcji użyteczności. Natomiast stosunkowo nowa metoda zaproponowana przez Abdellaoui, Bleichrodta i L'Haridona służąca do określania funkcji wartości na gruncie teorii perspektywy zostanie przedstawiona bardziej szczegółowo. Jej wykorzystanie będzie zilustrowane przykładem szacowania funkcji wartości pewnego decydenta. W przykładzie tym wykorzystany zostanie specjalnie skonstruowany przez autorkę arkusz ułatwiający przeprowadzenie interaktywnej procedury określania ekwiwalentów pewności przedstawianych loterii. W zakończeniu zostaną omówione zalety i wady wybranych metod.

W pracy wykorzystywany będzie zapis losowego wariantu decyzyjnego w postaci $(x, p; y)$, co należy rozumieć, że możliwe jest otrzymanie wyniku x z prawdopodobieństwem p lub wyniku y z prawdopodobieństwem $1 - p$.

1. Metody określania funkcji użyteczności

1.1. Metody ekwiwalentu pewności

W metodzie ekwiwalentu pewności (certainty equivalence method) w celu określenia funkcji użyteczności decydent jest proszony o określenie nielosowego wyniku w , zwanego ekwiwalentem pewności, dla którego zachodzi relacja indyferencji $(x, p; y) \sim w$, przy czym zakłada się, że przedział wartości X jest ciągły, a funkcja preferencji na przedziale wartości X jest ściśle rosnąca. Funkcję użyteczności określa się dla pewnego ustalonego przedziału wartości, przyjmując dolną x_0 i górną x^* jego granicę oraz $u(x_0) = 0$ i $u(x^*) = 1$. Stosowane są dwie wersje tej metody: frakcyjna oraz łańcuchowa.

We frakcyjnej metodzie ekwiwalentu pewności ustalany jest zbiór prawdopodobieństw $0 < p_1 < p_2 < \dots < p_n < 1$. Zadaniem decydenta jest określenie wartości ekwiwalentów pewności x_i w relacjach indyferencji o postaci $(x^*, p_i; x_0) \sim x_i$. Wyznaczenie użyteczności ekwiwalentu pewności jest proste, gdyż $u(x_i) = p_i$.

W łańcuchowej metodzie ekwiwalentu pewności rozpatruje się losowe warianty decyzyjne, w których oba wyniki są jednakowo prawdopodobne, czyli $p = 0,5$. W pierwszej kolejności decydent określa wartość x_1 , dla której zachodzi relacja indyferencji pomiędzy $(x^*, 0,5; x_0) \sim x_1$. Ze względu na przyjęte założenia $u(x_0) = 0$ i $u(x^*) = 1$ oraz $p = 0,5$, dla wartości x_1 otrzymujemy $u(x_1) = 0,5$. Następnie przedział wartości $[x_0, x^*]$ jest dzielony na dwa przedziały o postaci $[x_0, x_1]$ oraz $[x_1, x^*]$ i dla każdego z nich konstruowana jest odpowiednia loteria, a następnie decydent określa ich ekwiwalenty pewności, czyli wartości x_2 oraz odpowiednio x_3 , dla których zachodzą relacje indyferencji $(x_1, 0,5; x_0) \sim x_2$ oraz $(x^*, 0,5; x_1) \sim x_3$. Dla wartości x_2 oraz x_3 otrzymujemy $u(x_2) = 0,25$ oraz $u(x_3) = 0,75$. Uzyskane wartości x_2 oraz x_3 służą do dokonania następnych podziałów uzyskując przedziały o postaci $[x_0, x_2]$, $[x_2, x_1]$ oraz $[x_1, x_3]$, $[x_3, x^*]$, dla których konstruowane są odpowiednie loterie, a następnie decydent określa kolejne ekwiwalenty pewności. Postępowanie jest kontynuowane do momentu aż uzyska się zadowalającą liczbę wartości wystarczającą do estymacji funkcji użyteczności.

Dla obu wersji metody ekwiwalentu pewności, aby określić analityczną postać funkcji użyteczności, należy przyjąć jej parametryczną postać (charakteryzującą się koniecznymi i zaobserwowanymi własnościami) i na podstawie uzyskanych par wartości $(x_i, u(x_i))$ dokonać estymacji parametrów przyjętej funkcji użyteczności.

1.2. Metody ekwiwalentu prawdopodobieństwa

W metodzie ekwiwalentu prawdopodobieństwa (probability equivalence method) decydent określa prawdopodobieństwo p , dla którego zachodzi relacja indyferencji o postaci $(x, p; y) \sim w$, przy czym $x > w > y$. Funkcję użyteczności określa się dla wybranych wartości $x_0 < x_1 < \dots < x_{n+1}$, przyjmując $u(x_0) = 0$ i $u(x_{n+1}) = 1$. Stosowane są dwie wersje tej metody: ekstremalnych loterii oraz sąsiednich loterii.

W wersji ekstremalnych loterii decydent określa wartości prawdopodobieństw p_i , dla których zachodzi relacja indyferencji o postaci $(x_{n+1}, p_i; x_0) \sim x_i$. Ze względu na przyjęte założenia $u(x_0) = 0$ i $u(x_{n+1}) = 1$ wartości użyteczności dla kolejnych relacji wynoszą $u(x_i) = p_i$.

Wersja sąsiednich loterii jest znacznie trudniejsza. Decydent określa prawdopodobieństwa p_i , dla których zachodzą relacje indyferencji o postaci $(x_{i+1}, p_i; x_{i-1}) \sim x_i$. Zatem nie wykorzystuje się w przedstawianych loteriach wartości najlepszej i najgorszej, czyli x_{n+1} i x_0 , a zamiast tego wartości lokalnie najlepsze i najgorsze dla każdego x_i , czyli x_{i+1} i x_{i-1} . Na podstawie każdej odpowiedzi decydenta otrzymuje się zależności

$$u(x_i) = p_i u(x_{i+1}) + (1 - p_i) u(x_{i-1}) \quad (1)$$

Zakładając $u(x_0) = 0$, $u(x_{n+1}) = 1$, $f_0 = 2$ oraz $f_i = (1 - p_i)/p_i$ M. Novick i D. Lindley [9] rozwiązali układ n równań o postaci (1) z n niewiadomymi otrzymując

$$u(x_i) = \frac{\sum_{j=0}^{i-1} \prod_{k=0}^j f_k}{\sum_{j=0}^n \prod_{k=0}^j f_k} \quad \text{dla } i = 1, \dots, n \quad (2)$$

Zaletą tej wersji jest to, że wartości użyteczności nie są wrażliwe na precyzję oceny prawdopodobieństw, pod warunkiem, że decydent nie jest proszony o ocenę prawdopodobieństw bliskich 0 i 1.

Ostatnim etapem obu wersji metod jest estymacja parametrów analitycznej postaci przyjętej funkcji użyteczności.

W literaturze można znaleźć opisy metod badania funkcji użyteczności decydenta, które są znacznie rzadziej stosowane. Do metod tych należy np. metoda ekwiwalentu wartości (value equivalence method), hybrydowe metody oceny (hybrid assessment method), czy też metody oceny par loterii (paired-gamble methods). W metodach łączonych jednym z etapów estymacji funkcji

użyteczności jest ocena zgodności wyborów dokonywanych przez decydenta, co jest niezmiernie istotnym ich aspektem. Klasyfikację oraz krótkie opisy tych metod można znaleźć w pracy [6].

1.3. Metoda gamble-tradeoff

P. Wakker i D. Deneffe [11] zaproponowali metodę szacowania funkcji użyteczności wykorzystującą porównywanie loterii (gamble-tradeoff method) i można ją stosować zarówno do sytuacji, w których prawdopodobieństwa są znane, jak również do sytuacji decyzyjnych w warunkach niepewności. Dla uproszczenia rozpatrzmy sytuację decyzyjną w warunkach ryzyka, w której prawdopodobieństwa wystąpienia poszczególnych wyników są znane. Idea metody polega na określaniu wartości X , dla której decydent stwierdzi zachodzenie relacji indyferencji pomiędzy parą loterii o postaci: $(X, p; r)$ oraz $(x, p; R)$, przy czym $X > x$ oraz $R > r$, a wartości p, r, x oraz R są stałe. W podobny sposób określana jest wartość Y , dla której decydent jest indyferentny względem pary loterii o postaci $(Y, p; r)$ oraz $(y, p; R)$, przy czym $Y > y$ oraz $R > r$. Ponownie wartości p, r, x oraz R są stałe. Na gruncie teorii oczekiwanej użyteczności oraz z własności relacji indyferencji stwierdzamy, że zachodzą zależności

$$pu(X) + (1 - p)u(r) = pu(x) + (1 - p)u(R) \quad (3)$$

oraz

$$pu(Y) + (1 - p)u(r) = pu(y) + (1 - p)u(R) \quad (4)$$

Odejmując stronami równania otrzymujemy kolejno

$$\begin{aligned} pu(X) - pu(Y) &= pu(x) - pu(y) \\ u(X) - u(x) &= u(Y) - u(y) \end{aligned} \quad (5)$$

Zatem z analizy dwóch relacji indyferencji wnioskujemy o równości między różnicami użyteczności, co jest wykorzystywane do oszacowania funkcji użyteczności.

Pełna procedura określania funkcji użyteczności przebiega następująco. Eksperymentator przyjmuje dowolną, stałą wartość prawdopodobieństwa p oraz określa wartości referencyjne r i R takie, że $r < R$. Wartości te powinny być wystarczająco bliskie sobie, aby uzyskać wartości użyteczności z odpowiednim poziomem dokładności. Następnie określa minimalną wartość x_0 , dla której będzie określana użyteczność – może to być $x_0 = 0$. W kolejnym kroku konstruowane są loterie o postaci $(x_1, p; r)$ oraz $(x_0, p; R)$, przy czym wartość x_1 jest zmieniana do momentu, gdy decydent jest indyferentny wobec prezentowanej pary loterii. Następnie rozpatrywana jest para loterii o postaci $(x_2, p; r)$ oraz $(x_1, p; R)$, przy czym $x_2 > x_1$, a wartość x_2 jest zmieniana do momentu, gdy decydent jest indyferentny wobec drugiej pary loterii.

Analizując dwie relacje indyferencji eksperymentator otrzymuje zależność $u(x_2) - u(x_1) = u(x_1) - u(x_0)$. Przyjmując $u(x_0) = 0$ otrzymuje się $u(x_2) = 2u(x_1)$. Dla kolejnych wartości x_i , dla których decydent jest indyferentny względem par loterii $(x_i, p; r)$ oraz $(x_{i-1}, p; R)$, otrzymujemy $u(x_i) = i \cdot u(x_1)$. Postępowanie jest kontynuowane aż do uzyskania wystarczająco szerokiego zakresu wartości. Ponieważ użyteczność nie ma jednostek mianowanych, to za $u(x_1)$ można przyjąć dowolną dodatnią wartość np. $u(x_1) = 1/n$, gdzie n jest indeksem ostatniego otrzymanego wyniku x_n .

W powyższej procedurze wartości prawdopodobieństwa choć są jawnie podawane w loteriach, nie mają wpływu na obliczenia i określenie wartości użyteczności. Autorzy metody rozszerzają więc jej stosowanie także na sytuacje decyzyjne w warunkach niepewności, gdy prawdopodobieństwa zdarzeń są nieznane. Sytuacje takie są bliższe rzeczywistości. Rozpatrywane są losowe warianty decyzyjne zapisywane w postaci $(X, A; r)$, gdzie A jest zdarzeniem o nieznanym prawdopodobieństwie. Przykładem takiego losowego wariantu decyzyjnego może być „jeżeli operacja powiedzie się (zdarzenie A), to pacjent przeżyje X lat, jeżeli natomiast operacja nie powiedzie się, pacjent przeżyje r lat”. Pomimo niesprecyzowania wartości prawdopodobieństwa zdarzenia A wszystkie obliczenia i określanie wartości użyteczności są nadal poprawne.

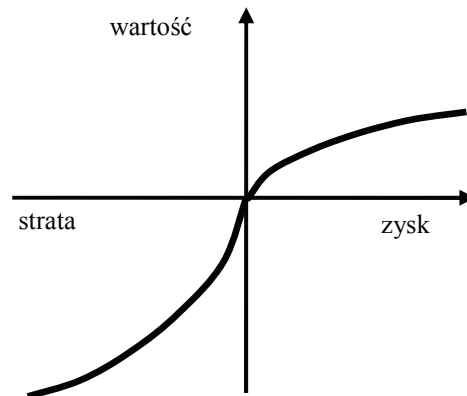
2. Szacowanie funkcji wartości

2.1. Własności funkcji wartości

Oceniając wartość wyniku decyzji porównują go do poprzedniego stanu nazywanego punktem referencyjnym. Funkcja wartości w teorii perspektywy [8] jest definiowana dla zmian (zysków i strat) w stosunku do tego punktu. Własności tej funkcji odzwierciedlają naturalne postrzeganie wartości. Wraz ze wzrostem zysków decydenci odczuwają co prawda coraz większą wartość, ale przyrosty tej wartości są coraz mniejsze. Podobne zjawisko ma miejsce w przypadku rozpatrywania strat – wraz ze wzrostem straty obniżenie wartości jej oceny jest coraz mniejsze. Zjawiska te są modelowane przez wypukłość funkcji wartości. W dziedzinie zysków funkcja ta jest wklęsła ($v''(x) < 0$ dla $x > 0$), a w dziedzinie strat jest wypukła ($v''(x) > 0$ dla $x < 0$), czyli krańcowe przyrosty tej funkcji zmniejszają się wraz ze wzrostem zysku lub straty.

Inna istotna własność funkcji wartości związana jest z odczuwaniem zmian poziomu bogactwa in plus i in minus. Z reguły decydenci bardziej odczuwają negatywne skutki straty niż przyjemność z zysku o tej samej wartości. Można powiedzieć, że „strata bardziej boli niż zysk o tej samej wartości cieszy”. Funkcja wartości charakteryzująca to zachowanie ma większy współczynnik nachylenia (jest bardziej stroma) dla strat niż dla zysków.

Funkcja wartości posiadająca przedstawione własności jest typu „S”. Przykładową funkcję przedstawiono na rys. 1. Postać analityczna tej funkcji wraz z oszacowaniami odpowiednich parametrów jest dobierana na podstawie preferencji decydentów ujawnionych w badaniach polegających na wskazywaniu preferowanego wariantu spośród dwóch proponowanych, przy czym w badaniu przedstawiana jest większa liczba par wariantów odpowiednio konstruowanych.



Rys. 1. Kształt funkcji wartości

Źródło: Opracowanie własne na podstawie [8].

Na podstawie przeprowadzanych eksperymentów w udziale studentów lub decydentów szacowane są postacie funkcji wartości. W pracy [10] zaproponowana została dwuczęściowa funkcja potęgowa postaci

$$v(x) = \begin{cases} x^\alpha, & x \geq 0 \\ -\lambda(-x)^\beta, & x < 0 \end{cases} \quad (6)$$

Wykorzystując procedurę nieliniowej regresji autorzy oszacowali parametry α , β i λ dla każdego uczestnika eksperymentu. Mediany tych wartości przyjęto jako oszacowania parametrów dla wszystkich uczestników badania. Parametry α i β wyniosły 0,88, natomiast parametr λ wyniósł 2,25. Wartości parametrów α i β wskazują na malejącą wrażliwość oceny na wzrost poziomu zysku lub straty, natomiast wartość parametru $\lambda > 1$ wskazuje na awersję do strat. Inne oszacowania parametrów funkcji wartości na podstawie przeprowadzanych eksperymentów można znaleźć np. w pracach [1; 4; 5].

2.2. Idea i etapy metody Abdellaoui-Bleichrodta-L'Haridona

Metoda zaproponowana przez M. Abdellaoui, H. Bleichrodta oraz O. L'Haridona [3] służy do określenia funkcji wartości na gruncie teorii perspektywy, łącząc zarówno podejście parametryczne, jak i nieparametryczne. Metoda ta może być wykorzystana także do określania funkcji wartości na gruncie innych teorii, w których w podobny sposób ocenia się warianty decyzyjne o dwóch możliwych wynikach. Teorie te to np. teoria rangowanej użyteczności (rank-dependent utility) czy też teoria awersji do niezadowolenia Gula (theory of disappointment aversion) [2; 7].

Do określenia funkcji wartości wykorzystywane jest podejście parametryczne, szeroko stosowane w literaturze przedmiotu. Główną zaletą przedstawianej metody jest brak założeń dotyczących funkcji ważenia prawdopodobieństw, gdyż wymagana jest znajomość wagi dla jednej wartości prawdopodobieństwa. Metoda opiera się na określaniu ekwiwalentu pewności dla 10-12 losowych wariantów decyzyjnych o dwóch niepewnych wynikach. Zaletą tej metody jest to, że nie jest ona podatna na przenoszenie błędów, gdyż rozpatrywane losowe warianty decyzyjne nie są ze sobą powiązane. Za pomocą tej metody można estymować parametry dowolnej funkcji wartości. Funkcja ta oczywiście musi spełniać założenia teorii perspektywy, czyli być rosnącą względem argumentu, z miejscem zerowym dla zerowej wartości argumentu, wypukłą w dziedzinie strat i wklęsłą w dziedzinie zysków. Funkcją najczęściej przyjmowaną w badaniach i najlepiej charakteryzującą odczucia decydentów jest funkcja potęgowa zaproponowana przez Tversky'ego i Kahnemana w pracy [10].

W procesie szacowania parametrów funkcji wartości wykorzystywane są binarne (o dwóch możliwych wynikach) losowe warianty decyzyjne, dla których teoria perspektywy i kumulacyjna teoria perspektywy są zbieżne i dają tę samą ocenę wariantu. W dalszej części przyjmuje się następujące określenia:

- wariant zysków to binarny losowy wariant decyzyjny nie zawierający ujemnych wyników, przy czym $x \geq y \geq 0$,
- wariant strat to binarny losowy wariant decyzyjny nie zawierający dodatnich wyników, przy czym $x \leq y \leq 0$,
- wariant mieszany to binarny losowy wariant decyzyjny zawierający jeden wynik ujemny i jeden dodatni, przy czym $x > 0 > y$.

Ocena wariantów na gruncie teorii perspektywy dokonywana jest zgodnie ze wzorami:

dla wariantu zysków

$$g^+(p)[v(x) - v(y)] + v(y) \quad (7)$$

dla wariantu strat

$$g^-(p)[v(x) - v(y)] + v(y) \quad (8)$$

dla wariantu mieszanego

$$g^+(p)v(x) + g^-(1-p)v(y) \quad (9)$$

gdzie:

$g^+(p)$ – funkcja ważenia prawdopodobieństw dla wyników dodatnich,

$g^-(p)$ – funkcja ważenia prawdopodobieństw dla wyników ujemnych,

$v(\cdot)$ – funkcja wartości.

Dodatkowo autorzy metody zakładają, że wartość wyników jest połączeniem współczynnika awersji do strat $\lambda > 0$, odzwierciedlającego odmienne odczuwanie zysków i strat, oraz samej wewnętrznej wartości wyniku. Zatem formalnie funkcja wartości jest zapisywana jako

$$v(x) = \begin{cases} V(x), & \text{gdy } x \geq 0 \\ \lambda V(x), & \text{gdy } x < 0 \end{cases} \quad (10)$$

Metoda określania funkcji wartości składa się z trzech etapów.

Etap 1. Określenie funkcji wartości w dziedzinie zysków.

Procedura rozpoczyna się przyjęciem pewnej wartości prawdopodobieństwa p_z stałej w tym etapie metody. Następnie eksperymentator definiuje serię złożoną z k wariantów zysków o postaci (x_i, p_z, y_i) , $i = 1, \dots, k$, przy czym $x_i \geq y_i \geq 0$.

Dla każdego tak zdefiniowanego wariantu decydent określa ekwiwalent pewności z_i . Wykorzystując wzory (7) oraz (10) otrzymujemy następującą zależność

$$v(z_i) = \delta^+[v(x_i) - v(y_i)] + v(y_i) \quad (11)$$

lub

$$z_i = v^{-1}(\delta^+[v(x_i) - v(y_i)] + v(y_i)) \quad (12)$$

gdzie z_i jest ekwiwalentem pewności i -tego binarnego losowego wariantu decyzyjnego (wariantu zysków) oraz $\delta^+ = g^+(p_z)$.

Korzyść wynikająca ze stałej wartości prawdopodobieństwa p_z polega na tym, że tylko jedna wartość funkcji ważenia prawdopodobieństwa jest wykorzystywana do oszacowania funkcji wartości, a zatem nie ma potrzeby dokonywania założeń odnośnie postaci funkcji ważenia prawdopodobieństw, a ponadto waga δ^+ może być potraktowana jako dodatkowy parametr, którego wartość należy oszacować.

W kolejnym kroku procedury należy przyjąć parametryczną postać funkcji wartości a następnie dokonać estymacji tej funkcji za pomocą nieliniowej metody najmniejszych kwadratów, przy czym parametrami, które należy oszacować są parametry funkcji wartości oraz waga δ^+ . Jeżeli przyjmiemy potęgową funkcję wartości $v(x) = x^\alpha$, należy oszacować parametry α oraz δ^+ następującego równania

$$z_i = \left(\delta^+ (x_i^\alpha - y_i^\alpha) + y_i^\alpha \right)^{1/\alpha} \quad (13)$$

Etap 2. Określenie funkcji wartości w dziedzinie strat.

Analogicznie przebiega procedura mająca na celu estymację funkcji wartości w dziedzinie strat. Najpierw eksperymentator zakłada pewną wartość prawdopodobieństwa $p_s = 1 - p_z$, a następnie określa serię złożoną z k wariantów strat o postaci (x_i, p_s, y_i) , $i = 1, \dots, k$, przy czym $x_i \leq y_i \leq 0$. Dla każdego tak zdefiniowanego wariantu decydent określa ekwiwalent pewności s_i .

Następnie wykorzystując wzory (8) oraz (10) otrzymujemy następującą zależność

$$v(s_i) = \delta^- [v(x_i) - v(y_i)] + v(y_i) \quad (14)$$

lub

$$s_i = v^{-1} \left(\delta^- [v(x_i) - v(y_i)] + v(y_i) \right) \quad (15)$$

gdzie s_i jest ekwiwalentem pewności i -tego wariantu strat oraz $\delta^- = g^-(p_s)$.

W ostatnim kroku tego etapu, przyjmując potęgową funkcję wartości (lub inną), należy oszacować za pomocą nieliniowej metody najmniejszych kwadratów, równanie o postaci

$$s_i = \left(\delta^- (x_i^\alpha - y_i^\alpha) + y_i^\alpha \right)^{1/\alpha} \quad (16)$$

Etap 3. Określenie współczynnika awersji do strat.

Określenie współczynnika awersji do strat λ następuje poprzez określenie kwoty straty w jednej relacji równoważności. W tym celu eksperymentator wybiera wartość z^* , która jest dowolną z wartości z_i uzyskanych w etapie 1. Zadaniem decydenta jest określenie takiej wartości straty s^* , dla której mieszany binarny losowy wariant decyzyjny o postaci: $(z^*, p_z; s^*)$ jest dla niego równoważny nielosowemu wynikowi 0. Z relacji równoważności wynika, że

$$\delta^+ \cdot v(z^*) + \delta^- \cdot \lambda v(s^*) = v(0) \equiv 0 \quad (17)$$

Ponieważ wielkości δ^+ , $v(z^*)$, δ^- , $v(s^*)$ są znane lub możliwe do wyliczenia z zależności (13) i (16) to wartość współczynnika awersji do strat można obliczyć następująco

$$\lambda = - \frac{\delta^+ \cdot v(z^*)}{\delta^- \cdot v(s^*)} \quad (18)$$

2.2. Funkcja wartości decydenta – przykład empiryczny

Empiryczne określanie funkcji wartości decydenta można przeprowadzić na podstawie ogólnej procedury zaproponowanej przez Farquhara [6]. Procedura ta składa się z pięciu kroków:

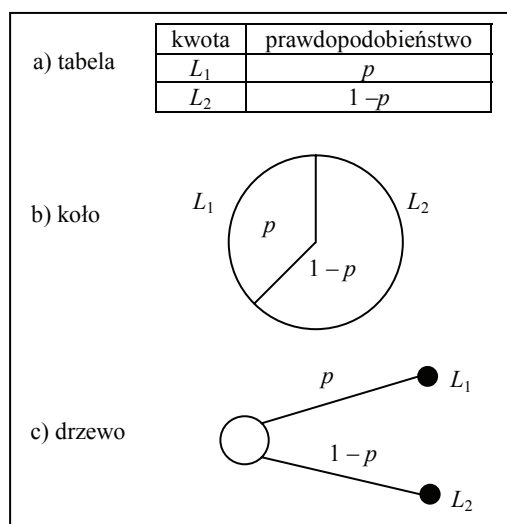
- Krok 1. Przygotowanie do oceny użyteczności polegające na określeniu struktury problemu decyzyjnego, przyjęciu odpowiedniej skali ocenianych wartości i zaznajomienie decydenta z różnymi aspektami procedury oceny.
- Krok 2. Zidentyfikowanie odpowiednich własności, takich jak monotoniczność, dziedzina, ciągłość lub własności ryzyka, co bezpośrednio wpływa na wybór dopuszczalnych postaci funkcji użyteczności.
- Krok 3. Określenie ilościowych ograniczeń nałożonych na funkcję użyteczności decydenta dokonywane poprzez porównywanie różnorodnych losowych wariantów decyzyjnych (loterii).
- Krok 4. Wybór ze zbioru parametrycznych funkcji użyteczności (spełniających własności zdefiniowane w kroku 2) tej, która spełnia wykryte ograniczenia ilościowe.
- Krok 5. Sprawdzenie spójności otrzymanych wyników i wniosków mające na celu wykrycie możliwych niespójności (i ich wyjaśnienie) i upewnienie, że otrzymane rezultaty opisują preferencje decydenta względem decyzji obarczonych ryzykiem.

Przedstawiona procedura proponuje pewną kolejność działań, jednakże bardzo często działania te są wykonywane w innej kolejności, a także możliwe jest powracanie do etapów już wykonanych, gdyż decydent w miarę przechodzenia przez kolejne etapy może zmieniać i udoskonalać swoje oceny. Odpowiedzi decydentów w początkowym etapie kroku 3 często ujawniają brak spójności preferencji. Dopiero później decydent zaczyna rozumieć wpływ wyrażonych wcześniej preferencji na kolejne warianty decyzyjne i zmienia wcześniejsze wskazania.

W dalszej części pracy przedstawiony zostanie przykład szacowania funkcji wartości pewnego decydenta dla wyników pomiędzy stratą 1000 zł a zyskiem 1000 zł za pomocą metody Abdellaoui, Bleichrodta i L'Haridona. W pierwszej kolejności zostaną skonstruowane loterie, następnie zadaniem decydenta będzie określenie ich ekwiwalentów pewności. Etap ten będzie wspomagany poprzez zadawanie dodatkowych pytań w formie par loterii, dla których decydent wskazuje loterię preferowaną. Po określeniu ekwiwalentów pewności przyjęta zostanie postać analityczna funkcji wartości i oszacowane zostaną jej parametry.

Sytuacja decyzyjna może być przedstawiona decydentowi w następujący sposób. Decydent stoi przed wyborem między dwoma możliwościami: pierwsza to otrzymanie kwoty C (nieobciążonej ryzykiem), natomiast druga to otrzymana-

nie biletu na loterię, w której można wygrać kwotę L_1 z prawdopodobieństwem p lub kwotę L_2 z prawdopodobieństwem $(1-p)$. Losowe warianty decyzyjne można przedstawić decydentowi za pomocą tabeli, koła lub drzewa (rys. 2), przy czym najbardziej sugestywna wydaje się forma koła. Zadaniem decydenta jest wskazanie wariantu preferowanego lub stwierdzenie indyferencji.



Rys. 2. Sposoby przedstawiania losowych wariantów decyzyjnych

Źródło: Opracowanie własne na podstawie [6].

W innym sformułowaniu problemu decydent może być proszony o bezpośrednie wskazanie kwoty C lub też jednej z kwot L_1 lub L_2 , lub też prawdopodobieństwa p . Ze względu na to, że decydentowi jest bardzo trudno bezpośrednio wskazać te wartości, w praktyce ich ustalenie sprowadza się do określania preferencji dla sekwencji loterii, w których odpowiednio zmieniana jest przedmiotowa wartość.

W celu oszacowania funkcji wartości w przedziale $[-1000, 1000]$ konstruowanych jest w pierwszym etapie 6 loterii z wynikami dodatnimi (zyskami) oraz w drugim etapie 6 loterii z wynikami ujemnymi (stratami). We wszystkich loteriach zakładamy, że oba wyniki są jednakowo prawdopodobne. Tabela 1 przedstawia przygotowane loterie nie zawierające strat (loterie Z1-Z6) oraz loterie nie zawierające zysków (S1-S6).

Tabela 1

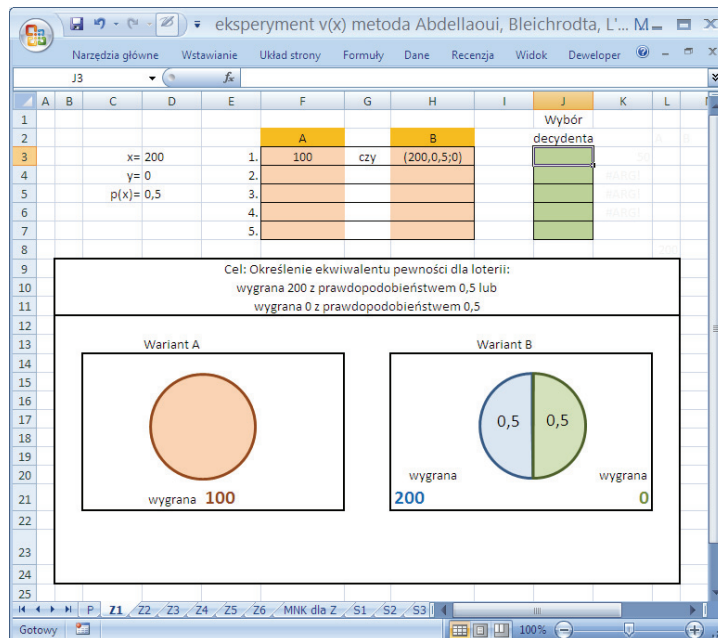
Loterie Z1-Z6 oraz S1-S6

Loterie nie zawierające strat						
Loteria	Z1	Z2	Z3	Z4	Z5	Z6
Wynik x	200	400	600	1000	1000	1000
Wynik y	0	0	0	0	600	800
Loterie nie zawierające zysków						
Loteria	S1	S2	S3	S4	S5	S6
Wynik x	-200	-400	-600	-1000	-1000	-1000
Wynik y	0	0	0	0	-600	-800

Dla każdej loterii decydent musi określić ekwiwalent pewności. Ponieważ w praktyce bardzo trudno jest podać kwotę równoważną loterii, to w sposób interaktywny można decydenta w tym wspomóc, prezentując kolejne pary wariantów decyzyjnych. W każdej parze jeden wariant (A) jest pewną kwotą (wygraną z prawdopodobieństwem 1), a drugi (B) – wyjściową loterią, której ekwiwalent ma określić decydent. Z dwóch zaprezentowanych wariantów decydent wskazuje preferowany. Kwota występująca w wariancie A jest początkowo wartością oczekiwaną wariantu B (loterii). Następnie, w zależności od wyboru wariantu preferowanego przez decydenta, kwota ta jest modyfikowana i ponownie decydent wskazuje preferowany wariant. Postępowanie to ma na celu przybliżenie wartości ekwiwalentu pewności i przeważnie po 5-6 porównaniach można otrzymać poszukiwaną wartość.

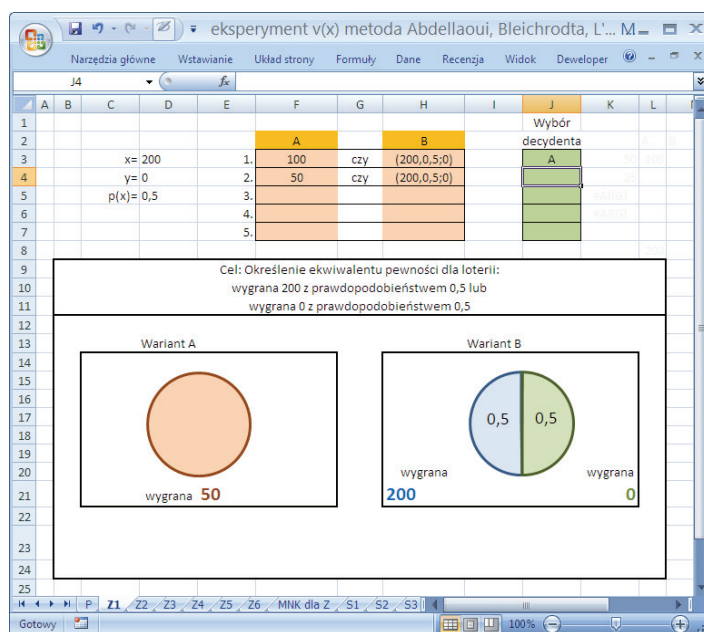
Modyfikacja kwoty realizowana jest w następujący sposób. Niech początkowo $G_0 = x$, $D_0 = y$ oraz $z_1 = G_0p + D_0(1 - p)$, przy czym x i y są wynikami wariantu ryzykownego (B) oraz $x > y$, a z_1 jest kwotą występującą w wariancie A w pierwszym porównaniu. Jeżeli w i -tym kroku decydent preferuje z przedstawionej pary wariant pozbawiony ryzyka (A), to $G_i = z_i$ i $D_i = D_{i-1}$. Jeżeli natomiast decydent preferuje wariant ryzykowny (B), to $D_i = z_i$ i $G_i = G_{i-1}$. Kwota przypisana wariantowi A w następnym kroku $i+1$ wynosi $z_{i+1} = G_i p + D_i(1 - p)$. Kwota obliczana po ostatnim wyborze decydenta jest jednocześnie oszacowaniem ekwiwalentu pewności wariantu ryzykownego.

W przykładzie dla loterii Z1 decydent ma określić ekwiwalent pewności. W celu prezentacji i usprawnienia dokonywania modyfikacji kwoty wariantu pozbawionego ryzyka utworzony został w Excelu specjalny arkusz. Po wprowadzeniu wartości x , y i p , definiujących wariant ryzykowny (loterię), pokazywane są oba warianty (rys. 3), z których decydent wybiera preferowany.

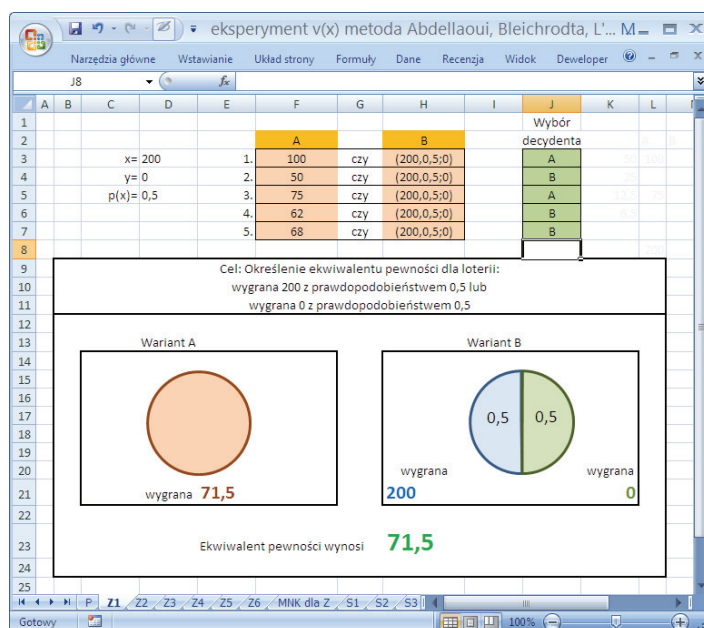


Rys. 3. Wybór przez decydenta preferowanego wariantu przy określaniu ekwiwalentu pewności dla loterii Z1 (pierwsze porównanie)

Wybory dokonywane przez decydenta wpisywane są w kolejne komórki w kolumnie J. Z pierwszej przedstawionej pary decydent wybrał wariant pozbawiony ryzyka A (komórka J3 na rys. 4). Po wpisaniu wyboru decydenta kwota występująca w wariantcie A jest modyfikowana zgodnie z procedurą przedstawioną powyżej. Następnie decydent ponownie proszony jest o wskazanie preferowanego wariantu dla kolejnych par. Rysunek 5 przedstawia kolejne wybory oraz wyznaczoną wartość ekwiwalentu pewności wynoszącą 71,5. W podobny sposób wyznaczane są ekwiwalenty pewności kolejnych loterii Z2-Z6 oraz S1-S6 (tabela 2).



Rys. 4. Wybór przez decydenta preferowanego wariantu przy określaniu ekwiwalentu pewności dla loterii Z1 (drugie porównanie)



Rys. 5. Ekwiwalent pewności dla loterii Z1

Tabela 2

Loterie Z1-Z6 oraz S1-S6 oraz wyznaczone ich ekwiwalenty pewności

Loterie nie zawierające strat						
Loteria	Z1	Z2	Z3	Z4	Z5	Z6
Wynik X	200	400	600	1000	1000	1000
Wynik Y	0	0	0	0	600	800
Ekwiwalent pewności	71,5	156	252,5	359	768,5	884
Loterie nie zawierające zysków						
Loteria	S1	S2	S3	S4	S5	S6
Wynik X	-200	-400	-600	-1000	-1000	-1000
Wynik Y	0	0	0	0	-600	-800
Ekwiwalent pewności	-84	-168,5	-252,5	-421,5	-781	-902,5

Na podstawie uzyskanych wartości ekwiwalentów pewności oraz samych loterii należy oszacować parametry funkcji wartości osobno w dziedzinie zysków i w dziedzinie strat. W przykładzie przyjęto funkcję potęgową jako funkcję wartości. Dla zysków estymacji parametrów funkcji o postaci (13) dokonano za pomocą regresji nieliniowej w programie PASW Statistics (rys. 6), przy czym zmienną zależną jest z , a zmiennymi niezależnymi x i y . Wartość parametru α oszacowano na poziomie 0,868, zatem funkcja wartości w dziedzinie zysków ma postać

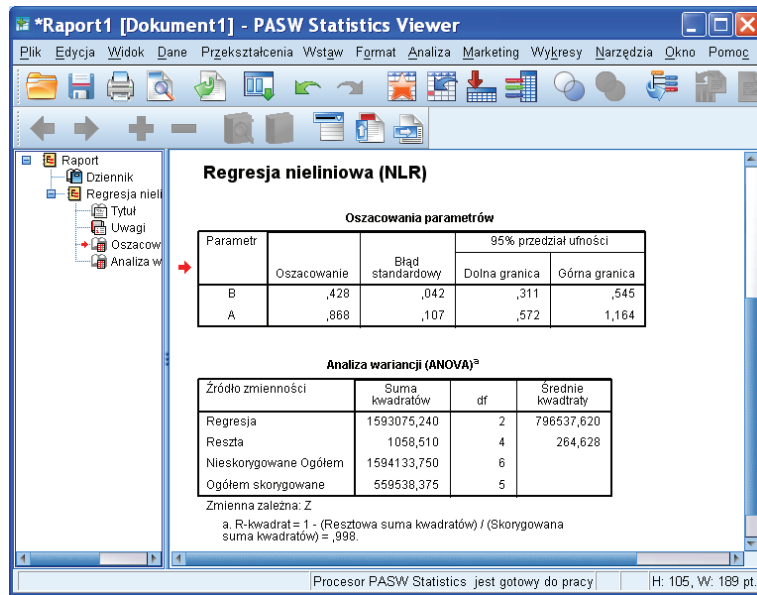
$$v(x) = x^{0,868} \quad \text{dla } x \geq 0 \quad (19)$$

Parametr $\delta^+ = g^+(0,5)$ przyjął wartość 0,428, co wskazuje, że ocena prawdopodobieństwa zysków jest przez decydenta zaniżana.

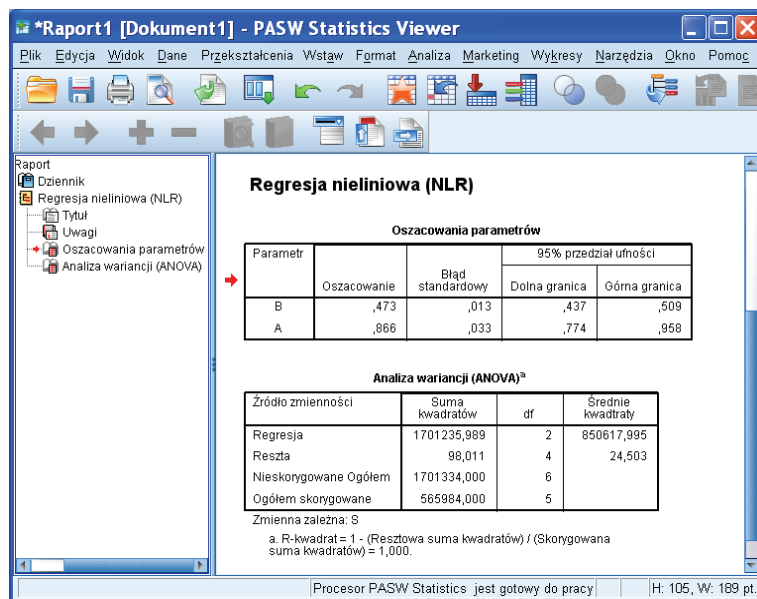
Podobnie dla loterii nie zawierających zysków oszacowano parametry zależności (16) za pomocą regresji nieliniowej (rys. 7). Parametr α ma wartość 0,866. Funkcja wartości w dziedzinie strat ma więc postać

$$v(x) = -(-x)^{0,866} \quad \text{dla } x < 0 \quad (20)$$

Natomiast wartość parametru $\delta^- = g^-(0,5) = 0,473$ wskazuje, że decydent zaniża ocenę prawdopodobieństwa strat.



Rys. 6. Oszacowanie parametrów funkcji wartości w dziedzinie zysków



Rys. 7. Oszacowanie parametrów funkcji wartości w dziedzinie strat

Ostatnia część eksperymentu ma na celu ustalenie współczynnika awersji do strat. Decydentowi przedstawiane są po dwa warianty, z czego jeden jest trywialny (wynik 0 z prawdopodobieństwem 1), a drugi umożliwia uzyskanie wartości równej ekwiwalentowi pewności dla loterii Z_i lub stratę. Celem tego postępowania jest ustalenie takiej wartości straty, dla której oba warianty są dla decydenta indyferentne. Ponieważ często jest trudno decydentowi zwerbalizować odpowiednią wartość straty to można go wspomóc przedstawiając serię par wariantów i prosząc o wskazanie preferowanego. Jako pierwsza para przedstawiany jest wariant trywialny oraz wariant $(Z, 0,5; S)$, gdzie wynik Z jest ekwiwalentem pewności loterii Z_1 , a $S = -Z_1^*$. Dla pierwszej pary wariantów przyjmujemy oznaczenia $S_1 = S$ oraz $S_0 = 0$. Jeżeli decydent wybierze wariant pozbawiony ryzyka (A) to strata występująca w wariantcie ryzykownym (B) jest zmieniana według wzoru

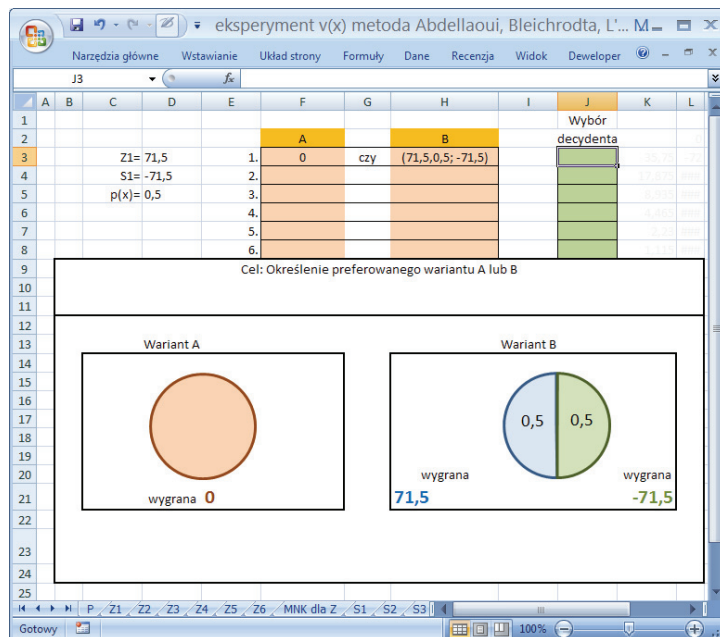
$$S_2 = S_1 + \frac{|S_1 - S_0|}{2}$$

Jeżeli natomiast decydent preferuje wariant B to strata S przyjmuje wartość

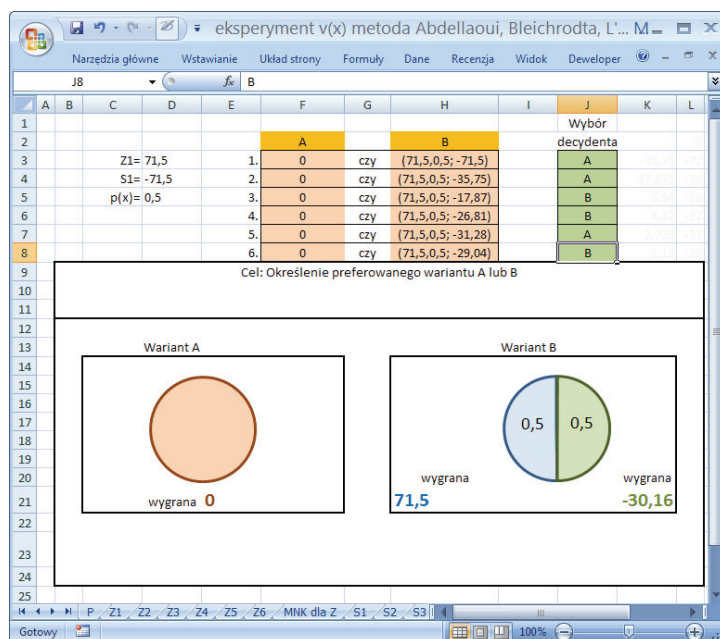
$$S_2 = S_1 - \frac{|S_1 - S_0|}{2}$$

Rysunek 8 przedstawia oba warianty początkowe występujące w analizowanym przykładzie. Decydent wybiera preferowany wariant i wskazuje wariant A. W kolejnej parze wariantów przedstawianych decydentowi strata wynosi $S_2 = -71,5 + \frac{|-71,5 - 0|}{2} = -35,75$. Decydent wskazuje kolejne preferowane warianty i w ten sposób po 6 porównaniach można podać przybliżoną wartość straty S , dla której oba warianty: ryzykowny (B) i pozbawiony ryzyka (A) są dla decydenta indyferentne. W przykładzie jest to wartość $-30,16$ (rys. 9). Podobne postępowanie przeprowadzane jest dla wartości ekwiwalentów pewności loterii Z_2 - Z_6 .

Na podstawie wartości strat wskazanych przez decydenta, wartości ekwiwalentów pewności dla loterii Z_1 - Z_6 oraz oszacowanych parametrów funkcji wartości w dziedzinie zysków i w dziedzinie strat możliwe jest oszacowanie współczynnika awersji do strat. Zależność pomiędzy podanymi wartościami przedstawiona jest wzorem (18), a tabela 3 zawiera dane i obliczenia.



Rys. 8. Wybór przez decydenta preferowanego wariantu w trzecim etapie metody (pierwsze porównanie)



Rys. 9. Określenie straty dla loterii mieszanej w trzecim etapie metody

Tabela 3

Wyznaczanie współczynnika awersji do strat

Loterie mieszane – s^* określone przez decydenta						
Loteria	M1	M2	M3	M4	M5	M6
Wynik z^*	71,5	156	252,5	359	768,5	884
Wynik s^*	-30,16	-46,31	-98,62	-106,57	-276,16	-207,18
λ	1,93	2,62	2,07	2,62	2,23	3,23

Dla loterii M1-M6 uzyskano wartości współczynnika awersji do strat z przedziału $\langle 1,93; 3,23 \rangle$, a średnia wartość wynosi 2,45. Wartość współczynnika większa od 1 wskazuje, że decydent ma awersję do strat, tzn. ponad dwa razy mocniej odczuwa stratę niż zysk o tej samej wartości bezwzględnej. Podsumowując, pełną postać oszacowanej funkcji wartości decydenta można zapisać następująco

$$v(x) = \begin{cases} x^{0,868} & \text{dla } x \geq 0 \\ -2,45(-x)^{0,866} & \text{dla } x < 0 \end{cases} \quad (21)$$

Otrzymana funkcja wartości spełnia założenia teorii perspektywy – jest rosnąca, wklęsła w dziedzinie zysków i wypukła w dziedzinie strat, z wyraźną awersją do strat.

Podsumowanie

Zarówno w teorii oczekiwanej użyteczności, jak i w teorii perspektywy, uzyskanie ocen decyzji zgodnych z preferencjami decydentów wymaga znajomości postaci funkcji użyteczności lub też funkcji wartości. W literaturze proponowane są różne metody wyznaczania parametrów tych funkcji. W zależności od grupy osób uczestniczących w eksperymentach uzyskiwane są różne wyniki, jednakże wyciągane na ich podstawie wnioski dotyczące głównych zachowań decydentów są przeważnie zbieżne.

W pracy dokonano przeglądu wybranych metod określania funkcji użyteczności oraz funkcji wartości. Podejście teoretyczne zostało wzbogacone przykładem praktycznego wykorzystania jednej z metod.

Niewątpliwą zaletą metod ekwiwalentu pewności i ekwiwalentu prawdopodobieństwa jest ich prostota i łatwość stosowania, natomiast głównymi wadami jest podatność na zniekształcenia wynikające z zależności między kolejnymi analizowanymi wariantami decyzyjnymi czy też zbyt szerokim przedziałem argumentów funkcji użyteczności.

W przypadku metody gamble-tradeoff zaletą jest brak podatności na efekt pewności oraz łatwe wykrywanie niekonsekwentnych odpowiedzi decydenta. Do wad należy zaliczyć podatność na przenoszenie błędów, większą czasochłonność oraz konieczność porównywania dwóch losowych wariantów decyzyjnych, zamiast losowego wariantu z nielosowym wynikiem.

Metoda wyznaczania funkcji wartości Abdellaoui, Bleichrodta i L'Hari-dona jest najbardziej pracochłonna, ale nie jest ona podatna na przenoszenie błędów (loterie są między sobą niezależne). Bardzo ważną jej zaletą jest też to, że nie trzeba dokonywać żadnych założeń odnośnie funkcji ważenia prawdopodobieństw.

Wszystkie procedury są bardzo pracochłonne, zarówno z punktu widzenia eksperymentatora, jak i decydenta. Aby określić funkcję użyteczności, wymagane jest skonstruowanie co najmniej 6-8 wariantów decyzyjnych, dla których decydent określa ekwiwalenty pewności lub prawdopodobieństwa. W praktyce decydent na początku eksperymentu nie jest świadomy swoich wyborów i ich wpływu na pozostałe warianty decyzyjne. Z tego względu często konieczny jest powrót do początkowych wariantów i zmiana wyborów. Drugim problemem jest samo podanie ekwiwalentu pewności. Aby pomóc decydentowi w trafnym wskazaniu tej wartości, dla każdego z 6-8 wariantów decyzyjnych zadawane jest 3-5 pytań pomocniczych. Łącznie tworzy to dość żmudny proces interaktywnego określania funkcji użyteczności.

Znacznie bardziej pracochłonny i żmudny jest proces określania funkcji wartości, w którym decydentowi przedstawiane jest 18 par wariantów. Tutaj także można pomóc decydentowi poprzez zadawanie kilku pytań pomocniczych, w których decydent ujawnia swoje preferencje. Zatem w tej procedurze decydent łącznie wskazuje prawie 100 swoich preferencji względem przedstawianych par wariantów.

Naukowcy starają się zaproponować nowe metody, w których będzie ograniczona liczba wariantów decyzyjnych przedstawiana decydentowi, a jednocześnie uzyskane wyniki jak najlepiej będą odzwierciedlać oceny decydenta dotyczące wartości, prawdopodobieństw, a także nastawienie do ryzyka. W pracach naukowych opisywane eksperymenty przeprowadzane są z reguły wśród studentów ekonomii, zarządzania, psychologii czy też medycyny. Aby zachęcić takich „decydentów” do współpracy, zastanowienia się nad przedstawianymi wariantami i ujawnienia własnych preferencji, konieczne jest stosowanie odpowiednich bodźców. Zachęta zwykle polega na zapłacie za udział w eksperymencie, a także na udziale jednej wylosowanej po eksperymencie osoby w rzeczywistej grze (z realną wygraną pieniężną) zgodnie z dokonanymi wcześniej wyborami.

Literatura

1. Abdellaoui M. (2000). Parameter-free Elicitation of Utilities and Probability Weighting Functions. *Management Science*, 46, s. 1497-1512.
2. Abdellaoui M., Bleichrodt H. (2007). Eliciting Gul's Theory of Disappointment Aversion by the Tradeoff Method. *Journal of Economic Psychology*, 28, s. 631-645.
3. Abdellaoui M., Bleichrodt H., L'Haridon O. (2008). A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory. *Journal of Risk and Uncertainty*, 36, s. 245-266.
4. Abdellaoui M., Vossman F., Weber M. (2005). Choice-based Elicitation and Decomposition of Decision Weights for Gains and Losses under Uncertainty. *Management Science*, 51(9), s. 1384-1399.
5. Donkers B., Melenberg B., van Soest A. (2001). Estimating Risk Attitudes Using Lotteries. A Large Sample Approach. *Journal of Risk and Uncertainty*, 22(2), s. 165-195.
6. Farquhar P. (1984). Utility Assessment Methods. *Management Science*, 30(11), s. 1283-1300.
7. Gul F. (1991). A Theory of Disappointment Aversion. *Econometrica*, 59(3), s. 667-686.
8. Kahneman D., Tversky A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47, s. 263-291.
9. Novick M., Lindley D. (1979). Fixed-state Assessment of Utility Functions. *Journal of American Statistical Association*, 74, s. 306-311.
10. Tversky A., Kahneman D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5, s. 297-323.
11. Wakker P., Deneffe D. (1996). Eliciting von Neumann-Morgenstern Utilities when Probabilities are Distorted or Unknown. *Management Science*, 42(8), s. 1131-1150.

SELECTED METHODS OF VALUATION OF UTILITY FUNCTION AND VALUE FUNCTION

Summary

In many studies utility function as well as value function and probability weighting function are assumed in the form estimated by their authors. These functions reflect personal decision-makers preferences and they ought to be estimated individually. This paper is a survey of selected methods of assessing utility function and value function. The certainty equivalence methods, probability equivalence methods and gamble-tradeoff method by P. Wakker and

D. Deneffe are presented. These methods are used for assessing utility function based on expected utility theory. The method of assessing value function (by Abdellaoui, Bleichrodt and L'Haridon) which is used to value uncertain decision alternative based on prospect theory is also presented in this paper. An example is an illustration of the last method.

Dorota Górecka

Michał Bernard Pietrzak

Uniwersytet Mikołaja Kopernika w Toruniu

ZASTOSOWANIE METODY PROMETHEE II W PROCESIE RANKINGOWANIA PROJEKTÓW EUROPEJSKICH W RAMACH REGIONALNEGO PROGRAMU OPERACYJNEGO WOJEWÓDZTWA KUJAWSKO-POMORSKIEGO NA LATA 2007-2013

Wprowadzenie

Europejska polityka regionalna^{*}, zwana też czasem polityką strukturalną, ma na celu wzmacnianie spójności społeczno-ekonomicznej Unii Europejskiej, niwelowanie różnic między poszczególnymi jej regionami i zapewnianie im wszechstronnego i harmonijnego rozwoju. Do instrumentów polityki regionalnej należą między innymi fundusze strukturalne^{**} oraz Fundusz Spójności. Ich zadaniem jest wspieranie restrukturyzacji i modernizacji gospodarek krajów członkowskich, finansowanie inwestycji infrastrukturalnych, promowanie zrównoważonego rozwoju, tworzenie nowych miejsc pracy, wzmacnianie konkurencyjności i innowacyjności regionów oraz wspieranie współpracy terytorialnej między nimi.

Unijna polityka regionalna w Polsce w obecnym okresie programowania wdrażana jest za pomocą pięciu programów operacyjnych (PO Infrastruktura i Środowisko, PO Kapitał Ludzki, PO Innowacyjna Gospodarka, PO Rozwój

^{*} W obecnym okresie programowania 2007-2013 polityka ta nazywana jest polityką spójności.

^{**} W poprzednim okresie programowania 2000-2006 funkcjonowały cztery fundusze strukturalne: Europejski Fundusz Rozwoju Regionalnego (EFRR), Europejski Fundusz Społeczny (EFS), Europejski Fundusz Orientacji i Gwarancji Rolnej (EFOiGR) oraz Finansowy Instrument Orientacji Rybołówstwa (FIOR). W obecnym okresie programowania 2007-2013 funkcjonują dwa: EFRR i EFS.

Polski Wschodniej, PO Pomoc Techniczna), Programów Europejskiej Współpracy Terytorialnej oraz szesnastu Regionalnych Programów Operacyjnych (po jednym dla każdego z województw).

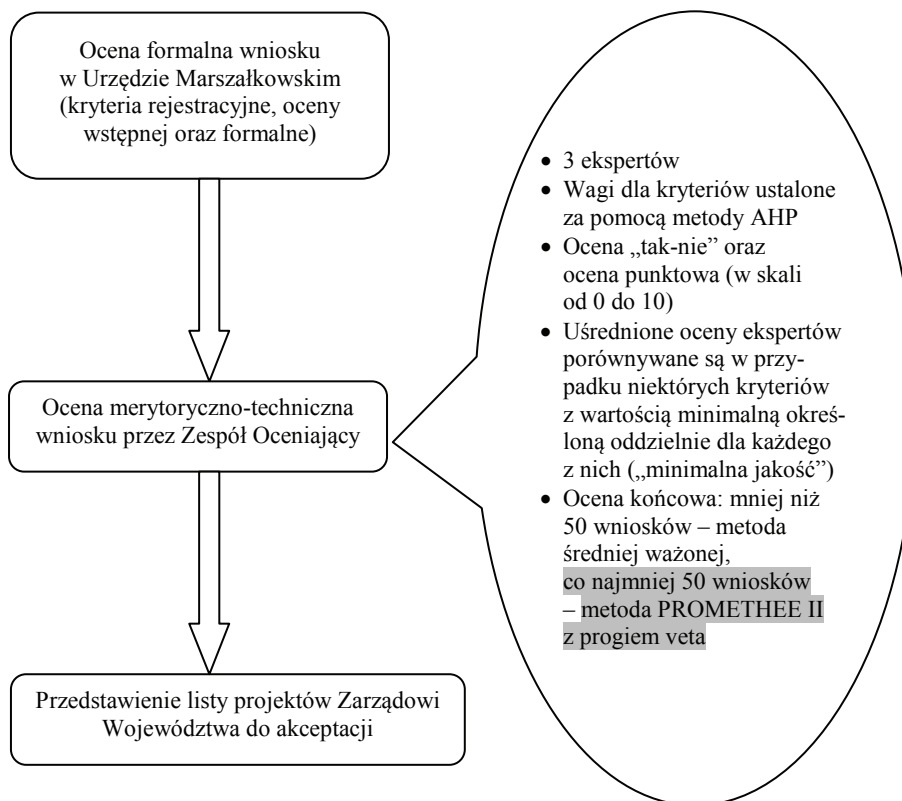
W ramach Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013 do rankingowania wniosków o dofinansowanie realizacji projektu z funduszy UE wykorzystywana jest między innymi metoda PROMETHEE II z progiem veta.

W niniejszym opracowaniu zaprezentowana zostanie aplikacja komputerowa opracowana na potrzeby systemu oceny projektów funkcjonującego w ramach RPO WK-P 2007-2013. Jej działanie przedstawione zostanie na przykładzie konkursu przeprowadzonego w latach 2008-2009 dla Poddziałania 5.2.2. *Wsparcie inwestycji przedsiębiorstw* realizowanego w ramach piątej osi priorytetowej RPO WK-P na lata 2007-2013 *Wzmocnienie konkurencyjności przedsiębiorstw*.

1. System wyboru projektów w ramach RPO WK-P na lata 2007-2013

Funkcjonujący w ramach RPO WK-P na lata 2007-2013 system oceny i wyboru projektów był kilkakrotnie modyfikowany, przy czym w latach 2008-2009, kiedy to przeprowadzany był konkurs w ramach Poddziałania 5.2.2. *Wsparcie inwestycji przedsiębiorstw*, przedstawiał się on jak na rys. 1.

Celem Poddziałania 5.2.2. *Wsparcie inwestycji przedsiębiorstw* Działania 5.2. RPO WK-P 2007-2013 o tej samej nazwie jest podniesienie konkurencyjności przedsiębiorstw (również tych działających w branży turystycznej) poprzez ich unowocześnianie oraz zwiększanie efektywności ich działania skutkujące wzrostem zatrudnienia [9, s. 111].



Rys. 1. Procedura oceny i wyboru wniosków o dofinansowanie realizacji projektu z funduszy Unii Europejskiej obowiązująca w 2009 roku w ramach Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013

Podczas zakończonego w 2009 roku konkursu dla tego Poddziałania* każdy wniosek o dofinansowanie realizacji projektu oceniany był pod względem merytoryczno-technicznym przez złożoną z 3 ekspertów Komisję Konkursową, która brała pod uwagę następujące kryteria**:

* 3 sierpnia 2009 roku Zarząd Województwa Kujawsko-Pomorskiego uchwałą nr 60/1014/09 zatwierdził do dofinansowania 97 projektów złożonych na konkurs z realizowanego w ramach RPO WK-P na lata 2007-2013 Poddziałania 5.2.2. *Wsparcie inwestycji przedsiębiorstw*. Był to pierwszy konkurs, w którym rankingowanie przeprowadzone zostało z wykorzystaniem metody PROMETHEE II. <http://www.mojregion.eu/regionalny-program-operacyjny-wojewodztwa-kujawsko-pomorskiego/konkursy/wyniki/os-priorytetowa-5/czytaj/items/dofinansowanie-dla-przedsiębiorców.965.html> (10.03.2011).

** Patrz: aneks.

Tabela 1

Kryteria merytoryczno-technicznej dopuszczalności projektu

Nr	Kryterium	Ocena	
B.1.1	Cele projektu wspierają realizację celów określonych dla Działania 5.2.	TAK	NIE
B.1.2	Wykonalność techniczna projektu	TAK	NIE
B.1.3	Trwałość rezultatów projektu	TAK	NIE
B.1.4	Wykonalność finansowa projektu	TAK	NIE

Źródło: [10, s. 88-89].

Tabela 2

Kryteria oceny jakości projektu

Nr	Kryterium	Ocena
B.2.1	Oczekiwane rezultaty projektu w sposób bezpośredni wspierają realizację celów określonych w Uszczegółowieniu Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego dla Działania 5.2. poprzez poprawę oferty produktowej i technologicznej przedsiębiorstwa	0-10
B.2.2	Projekt jest zgodny z preferowanymi typami projektów	0-10
B.2.3	Oczekiwane rezultaty projektu w sposób bezpośredni wspierają realizację celów określonych w Uszczegółowieniu Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego dla Działania 5.2. poprzez zwiększenie efektywności działania przedsiębiorstwa	0-10
B.2.4	Oczekiwane rezultaty projektu w sposób bezpośredni wspierają realizację celów określonych w Uszczegółowieniu Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego dla Działania 5.2. poprzez zwiększenie konkurencyjności działania przedsiębiorstwa	0-10
B.2.5	Udział środków własnych beneficjenta w projekcie	0-10
B.2.6	Wypełnienie polityk i zasad wspólnotowych	0-10
B.2.7	Strategiczny charakter projektu i komplementarność z innymi przedsięwzięciami	0-10
B.2.8	Promocja projektu	0-10

Źródło: [10, s. 89-91].

Tabela 3

Kryteria oceny wpływu przewidywanych rezultatów projektu na osiągnięcie założonych w RPO celów

Nr	Kryterium	Ocena
B.3.1	Wpływ projektu na rynek pracy	0-10

Źródło: [10, s. 91].

2. Metoda PROMETHEE II z progiem veta

Metody z rodziny PROMETHEE* to grupa technik wielokryterialnych należących do klasy metod przewyższania, opartych na metodologii zaproponowanej przez Roy [7].

Model preferencji w metodach z rodziny PROMETHEE powstaje na podstawie porównania wariantów decyzyjnych (projektów) parami ze sobą, przy czym pod uwagę brane są różnice między ocenami tych wariantów (projektów) dla wszystkich rozpatrywanych kryteriów. Im większa jest różnica pomiędzy ocenami, tym silniej – ze względu na dane kryterium – preferowany jest jeden z wariantów (projektów). Jeśli różnica ta jest niewielka, to jeden z wariantów (projektów) jest słabo preferowany względem drugiego lub też warianty (projekty) uznawane są za równoważne ze względu na to, że występująca pomiędzy ich ocenami różnica jest nieistotna (mało znacząca), jeśli natomiast różnica w ocenach jest dostatecznie duża, to jeden z wariantów (projektów) jest silnie preferowany ze względu na dane kryterium. Rozróżnienie sytuacji preferencyjnych jest możliwe dzięki wprowadzeniu do analizy progów równoważności i preferencji związanych z poszczególnymi kryteriami.

W zastosowanym w ramach RPO WK-P na lata 2007-2013 podejściu przyjęte zostało ponadto założenie o ograniczonej kompensacji. Oznacza to, że możliwa jest taka sytuacja, w której przewaga wariantu (projektu) *a* nad wariantem (projektem) *b* ze względu na jedno z kryteriów powoduje, że wariant (projekt) *b* nie powinien zostać uznany za lepszy od wariantu (projektu) *a* nawet wtedy, gdy jest on preferowany ze względu na wszystkie pozostałe kryteria. Do analizy tego typu sytuacji wykorzystywany jest próg veta.

Ostateczny wynik danego wariantu (projektu) stanowi różnicę pomiędzy dodatnim a ujemnym przepływem przewyższania. Jeśli różnica ta jest większa od zera, to oznacza to, że dany wariant (projekt) bardziej przewyższa inne warianty (projekty) ze względu na cały zestaw rozpatrywanych kryteriów, niż one przewyższają jego, natomiast w wypadku gdy różnica ta jest mniejsza od zera, dany wariant (projekt) jest w większym stopniu przewyższany przez inne warianty (projekty), niż sam je przewyższa. Im wyższa jest wartość różnicy pomiędzy przepływami przewyższania, tym dany wariant decyzyjny (projekt) jest lepszy.

* Akronim PROMETHEE pochodzi od słów: Preference Ranking Organisation METHod for Enrichment Evaluations (metoda konstrukcji rankingu preferencji w celu wzbogacenia ewaluacji). Prezentację metodologii i technik PROMETHEE można znaleźć w pracach [1; 2; 3].

Schemat postępowania w metodzie PROMETHEE II z progiem veta jest zatem następujący:

1. Obliczenie wartości współczynników zgodności $c(a_i, a_j)$ dla wszystkich par wariantów (a_i, a_j) :

$$c(a_i, a_j) = \sum_{k=1}^n w_k \varphi_k(a_i, a_j),$$

gdzie:

$$\sum_{k=1}^n w_k = 1,$$

$$\varphi_k(a_i, a_j) = \begin{cases} 1, & \text{gdy } f_k(a_i) - f_k(a_j) > p_k, \\ \frac{f_k(a_i) - f_k(a_j) - q_k}{p_k - q_k}, & \text{gdy } q_k < f_k(a_i) - f_k(a_j) \leq p_k, \\ 0 & \text{w innych przypadkach,} \end{cases}$$

przy czym:

- w_k – waga przypisana k -temu kryterium,
- $f_k(a_i)$ – ocena obiektu a_i względem kryterium f_k ,
- q_k – próg równoważności* dla kryterium f_k ,
- p_k – próg preferencji** dla kryterium f_k .

2. Obliczenie wskaźników niezgodności $d_k(a_i, a_j)$ dla każdej pary wariantów względem każdego z kryteriów:

$$d_k(a_i, a_j) = \begin{cases} 1, & \text{gdy } f_k(a_j) - f_k(a_i) > v_k, \\ \frac{f_k(a_j) - f_k(a_i) - p_k}{v_k - p_k}, & \text{gdy } p_k < f_k(a_j) - f_k(a_i) \leq v_k, \\ 0 & \text{w innych przypadkach,} \end{cases}$$

gdzie:

- v_k – próg veta*** dla kryterium f_k .

* Próg równoważności – maksymalna różnica pomiędzy ocenami dwóch obiektów, przy której nie można jeszcze stwierdzić, że jeden z nich jest lepszy (pod względem danego kryterium), w związku z czym obiekty te będą traktowane jako równoważne (z punktu widzenia danego kryterium).

** Próg preferencji – minimalna różnica pomiędzy ocenami dwóch obiektów, po przekroczeniu której nie istnieją już żadne wątpliwości, że jeden z nich jest lepszy (pod względem danego kryterium).

*** Próg veta – minimalna różnica pomiędzy ocenami dwóch obiektów dla danego kryterium, po przekroczeniu której wykluczone jest uznanie ocenionego niżej obiektu za ogólnie lepszy od tego ocenionego pod względem danego kryterium wyżej, nawet jeśli zostanie on uznany za zdecydowanie lepszy od drugiego pod względem wszystkich innych kryteriów.

3. Obliczenie współczynników wiarygodności $\sigma(a_i, a_j)$:

$$\sigma(a_i, a_j) = c(a_i, a_j) \prod_{k \in D(a_i, a_j)} \frac{1 - d_k(a_i, a_j)}{1 - c(a_i, a_j)},$$

gdzie: $D(a_i, a_j) = \{k : d_k(a_i, a_j) > c(a_i, a_j)\}.$

Współczynnik wiarygodności jest co najwyżej równy wartości współczynnika zgodności $c(a_i, a_j)$ w sytuacji, gdy nie ma opozycji i jest tym mniejszy, im jest ona silniejsza; w szczególności gdy istnieje kryterium, które stawia veto ($d_k(a_i, a_j) = 1$), to współczynnik wiarygodności wynosi 0.

Rezultatem przedstawionych obliczeń jest macierz współczynników wiarygodności.

4. Wyznaczenie wartości przepływu przewyższania netto dla każdego wariantu:

$$\phi(a_i) = \phi^+(a_i) - \phi^-(a_i),$$

gdzie:

$$\phi^+(a_i) = \sum_{j=1}^m \sigma(a_i, a_j),$$

$$\phi^-(a_i) = \sum_{j=1}^m \sigma(a_j, a_i).$$

5. Skonstruowanie finalnego rankingu wariantów na podstawie uporządkowanych malejąco wartości przepływów netto (wariant a_i jest lepszy od wariantu a_j , jeśli $\phi(a_i) > \phi(a_j)$; jeśli zaś $\phi(a_i) = \phi(a_j)$, to warianty a_i i a_j są równoważne).

3. Opis aplikacji komputerowej

Prezentowana w opracowaniu aplikacja komputerowa stworzona została na zamówienie Urzędu Marszałkowskiego Województwa Kujawsko-Pomorskiego. Podstawową funkcją aplikacji miała być obsługa systemu oceny projektów w ramach ogłoszonego w 2008 roku, a zakończonego w roku 2009 konkursu dla Poddziałania 5.2.2. *Wsparcie inwestycji przedsiębiorstw* realizowanego w ramach piątej osi priorytetowej RPO WK-P na lata 2007-2013 *Wzmocnienie konkurencyjności przedsiębiorstw*. Program posiada graficzny interfejs, którego strona główna przedstawiona została na rys. 2.

Urząd Marszałkowski Województwa Kujawsko-Pomorskiego

Regionalny Program Operacyjny Województwa Kujawsko-Pomorskiego

Działanie 5.2, poddziałanie 5.2.2 Wsparcie inwestycji przedsiębiorstw

Strona główna programu [Instrukcja obsługi programu](#)

Aktualnie zadeklarowana liczba ocenianych wniosków [Określ/zmień liczbę wniosków](#)

Jeżeli liczba ocenianych wniosków została wprowadzona i jest poprawna, przejdź do wprowadzania danych

[Wprowadzanie danych](#)

[Ranking wniosków](#)

[Generowanie list wniosków](#)

[Zakończ pracę z programem i zapisz](#)

Rys. 2. Strona główna programu

Wszystkie strony programu posiadają wymaganą przez zamawiającego instrukcję obsługi. Strona główna pozwala na przejście do kolejnych stron, pełniących specjalistyczne funkcje. Możliwe jest tutaj ustalenie liczby wszystkich ocenianych wniosków za pomocą przycisku „Określ/zmień liczbę wniosków”. Przycisk „Wprowadzanie danych” pozwala na przejście do strony, na której dla każdego wniosku wprowadzane są niezbędne do przeprowadzenia procedury rankingowania dane. Następną opcję – wykonania rankingu wniosków – gwarantuje przycisk „Ranking wniosków”. Do porządkowania wniosków wykorzystywana jest metoda PROMETHEE II z progiem veta. Kolejną dostępną funkcją jest generowanie list projektów o różnym statusie, co zapewnia przycisk „Generowanie list wniosków”. Ostatnim przyciskiem na stronie głównej jest przycisk pozwalający na zapisanie wyników i zakończenie pracy z programem.

Przycisk „Wprowadzanie danych” otwiera stronę przedstawioną na rys. 3. Zaprezentowana strona umożliwia wprowadzanie danych dotyczących wniosku oraz ocen każdego z trzech ekspertów. Do zespołu oceniającego należą przewodniczący powołany z ramienia Urzędu, ekspert ekonomiczny oraz ekspert techniczny. Zatwierdzone dane zapisywane są w bazie, a użytkownik – wybierając właściwy numer wniosku – może w każdej chwili do nich dotrzeć

i dokonać w nich niezbędnych zmian. W ramach przeprowadzanej oceny wnioski może posiadać status wniosku rozpatrzonego pozytywnie, rozpatrzonego negatywnie, wniosku cofniętego do ponownej oceny formalnej bądź też przekazanego do wyjaśnienia. Taki stan rzeczy prowadzi do tego, że każdy wniosek może zostać poddany ocenie czterokrotnie. Problem ten rozwiązany został w omawianym programie poprzez wprowadzenie czterech zakładki („Etap oceny”); w każdej z nich można od nowa wpisać dane dotyczące oceny wniosku. Podczas omawianego konkursu dane w większości przypadków wprowadzane były jedynie w zakładce „Etap I”, ponieważ złożone wnioski od razu uzyskiwały status rozpatrzonych pozytywnie albo negatywnie. Tylko wnioski rozpatrzone pozytywnie poddawane są procedurze rankingowania. Istnieje możliwość porządkowania wniosków w różnych rankingach, co uzyskiwane jest poprzez przypisanie każdemu wnioskowi numeru rankingowego. Pozostałe wnioski, którym przypisuje się wartość zero w polu „Numer rankingowy”, nie są rankingowane.

Urząd Marszałkowski Województwa Kujawsko-Pomorskiego

Liczba ocenianych wniosków: 10 | Wybierz liczbę porządkową wniosku: 10 | Kasowanie wprowadzonego wniosku lub etapu | Instrukcja obsługi programu | Kasowanie wniosku | Kasowanie etapu | Powrót do strony głównej

Etapy wprowadzania wniosków | **Wprowadzanie danych** | Numer protokołu oraz numer wniosku | **Znajdź wniosek**

Etap I | Etap II | Etap III | Etap IV

Data przeprowadzonej oceny (dd.mm.rrrr): sdf | Protokół: d | RPKP-05.02.02-04- | 10 | /08

Ocena wniosku

☒ Ocena pozytywna
☐ Ocena negatywna
☐ Ponowna ocena formalna
☐ Do wyjaśnienia

Nazwa wnioskodawcy: sdf | Tytuł projektu: sdf | Lokalizacja: sdf

Numer rankingowy: 1 | Całkowita wartość wniosku: 700 | Wartość wnioskowanego dofinansowania: 450

Eksperti

	Przewodniczący	Ekonomiczny	Techniczny
Nazwisko	sdf	df	df
Imię, imiona	sdf	sdf	sdf
PESEL	7777777777	3333333333	5555555555

Kryteria oceny

	Przewodniczący	Ekonomiczny	Techniczny	Srednie
Kryterium B 1.1	tak	tak	tak	
Kryterium B 1.2	tak	tak	tak	
Kryterium B 1.3	tak	tak	tak	
Kryterium B 1.4	tak	tak	tak	
Kryterium B 2.1	9	8	10	9
Kryterium B 2.2	7	8	6	7
Kryterium B 2.3	7	5	8	6,7
Kryterium B 2.4	8	8	8	8
Kryterium B 2.5	5	5	6	5,3
Kryterium B 2.6	6	5	6	5,7
Kryterium B 2.7	8	8	7	7,7
Kryterium B 2.8	7	5	8	6,7
Kryterium B 3.1	7	8	8	7,7

Drukowanie protokołów

Przewodniczący | Drukuj protokół
Ekonomiczny | Drukuj protokół
Techniczny | Drukuj protokół
Ocena zbiorcza | Drukuj protokół

Zapisywanie protokołów

Przewodniczący | Zapisz protokół
Ekonomiczny | Zapisz protokół
Techniczny | Zapisz protokół
Ocena zbiorcza | Zapisz protokół

Akceptuj wprowadzone dane lub dokonane zmiany

Rys. 3. Strona pozwalająca na wprowadzanie danych

Przycisk „Kasowanie wniosku” pozwala na usunięcie wybranego wniosku z bazy w całości, natomiast przycisk „Kasowanie etapu” umożliwia usunięcie danych dotyczących tylko jednego, wybranego etapu. Dodatkowo strona pozwala na drukowanie protokołów oceny wniosku dla każdego z ekspertów, a także na wydruk protokołu zbiorczego. Istnieje również możliwość zapisu protokołów do pliku programu Excel. Procedurę wprowadzania danych oraz drukowania protokołów kończy przejście do strony głównej.

W przypadku wprowadzenia ocen wszystkich zadeklarowanych wniosków program udostępnia funkcję ich rankingowania. Strona dotycząca sporządzania rankingów przedstawiona została na rys. 4.

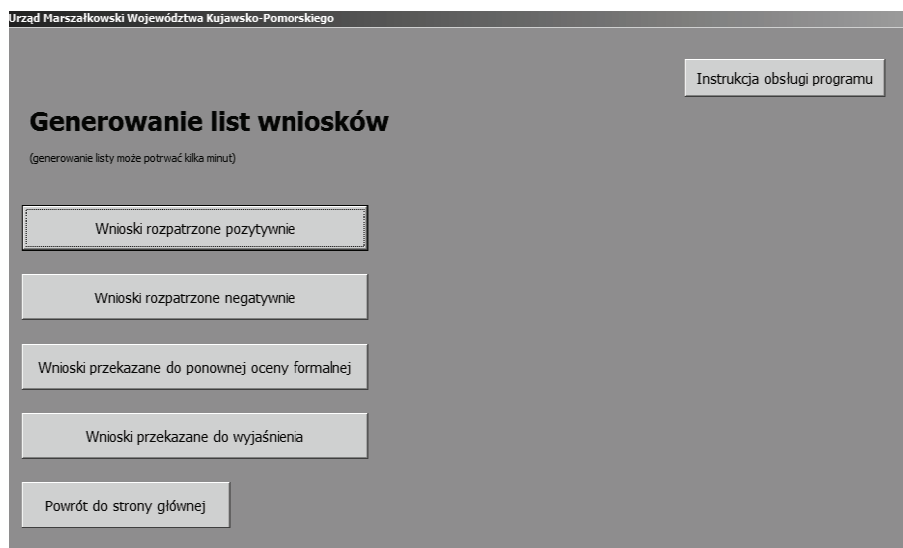


Rys. 4. Strona umożliwiająca tworzenie rankingów wniosków

Na stronie tej dostępne są dwie opcje. Pierwsza z nich polega na przeprowadzeniu – po uprzednim zadeklarowaniu numeru rankingu – porządkowania wniosków i utworzeniu ich listy, która następnie może zostać wydrukowana. Druga opcja pozwala na sporządzenie rankingu i zapisanie wyników do pliku. Lista rankingowa, oprócz zajętych przez poszczególne projekty miejsc, zawiera również podstawowe dane dotyczące wniosków, a także skumulowane wartości dofinansowania. Przy znanej wielkości dostępnych środków pozwalają one na podjęcie decyzji, ile wniosków otrzyma dofinansowanie. Po przekroczeniu planowanego budżetu kolejne wnioski z listy rankingowej otrzymują status niezakwalifikowanych do dofinansowania.

Następną funkcją prezentowanej aplikacji komputerowej jest tworzenie list wniosków o wybranym statusie. Wybranie na stronie głównej przycisku „Generowanie list wniosków” pozwala na przejście do strony generowania list wniosków. Dostępne są na niej listy wniosków rozpatrzonych pozytywnie

i negatywnie, listy wniosków skierowanych do ponownej oceny formalnej oraz tych przekazanych do wyjaśnienia. Funkcja tworzenia i drukowania wybranych list pozwala na prosty podział wniosków, a także umożliwia komunikację z potencjalnymi beneficjentami środków unijnych.



Rys. 5. Strona umożliwiająca tworzenie różnych zestawień wniosków

Ostatnią funkcją programu jest zapisywanie jego aktualnej wersji. Wybór przycisku „Zakończ pracę z programem i zapisz” powoduje zapisanie aktualnej wersji programu pod nazwą „ocena” w katalogu „baza”. Aplikacja jest każdorazowo wywoływana poprzez uruchomienie tego pliku. Dodatkowo program zapisywany jest w katalogu „Archiwum” pod nazwą „ocena” rozszerzoną o datę i godzinę zapisu. Plik ten daje możliwość powrotu do historycznych wersji programu.

Prezentowany program został wykorzystany do obsługi wymienionego wyżej konkursu. Pracownicy Urzędu Marszałkowskiego Województwa Kujawsko-Pomorskiego stosowali wszystkie dostępne funkcje zawarte w aplikacji, co pozwoliło na sprawne i profesjonalne przeprowadzenie procedury konkursowej.

4. Korzyści płynące z wykorzystania w procesie wyboru projektów metody PROMETHEE II z progiem veta

Wykorzystanie w procesie ewaluacji i selekcji projektów europejskich przyjaznej dla swoich użytkowników metody PROMETHEE II z progiem veta pozwala nie tylko na ograniczenie możliwości manipulowania wynikami, ale także na urealnienie analizy preferencji, gdyż technika ta – dzięki zastosowaniu progu równoważności i progu preferencji – uwzględnia fakt, iż realnym problemom decyzyjnym towarzyszy niepewność, nieprecyzyjność, nieokreśloność i zmienność dotycząca zarówno danych, jak również ocen oraz preferencji uczestników procesu decyzyjnego, którzy oceniając nierzadko nawet kilkadziesiąt projektów w ciągu kilku dni nie muszą być wcale do końca konsekwentni w swoich działaniach. Z kolei zastosowanie progu veta wyklucza możliwość pełnej kompensacji, w związku z czym możliwa jest taka sytuacja, w której przewaga projektu a_i nad projektem a_j ze względu na jedno z kryteriów jest na tyle duża, że wyklucza uznanie projektu a_j za lepszy od projektu a_i nawet wtedy, gdy jest on preferowany ze względu na wszystkie pozostałe kryteria [6, s. 103-104]. Pełna kompensacja jest zjawiskiem szczególnie niepożądanym w przypadku kryteriów związanych z wykonalnością techniczną projektu oraz doбором technologii, przygotowaniem organizacyjno-instytucjonalnym i finansowym beneficjenta, a także oddziaływaniem projektu na środowisko naturalne; budzić też może poważne wątpliwości w przypadku kryteriów związanych z realizacją celów działania czy poddziałania, w ramach którego dany projekt ubiega się o dofinansowanie [4, s. 236-237].

Aby dofinansowywane w ramach RPO WK-P 2007-2013 projekty posiadały zadowalającą jakość, oprócz progów równoważności, preferencji i veta dla wybranych kryteriów określone zostały konieczne do uzyskania minimalne wartości średnich arytmetycznych z ocen ekspertów.

W celu pokazania opisanych wyżej korzyści płynących z zastosowania metody PROMETHEE II z progiem veta technika ta wykorzystana została w symulacji procesu rankingowania projektów przeprowadzonej z wykorzystaniem opracowanego na potrzeby Urzędu Marszałkowskiego programu komputerowego. Pod uwagę wzięto 10 fikcyjnych wniosków aplikacyjnych o dofinansowanie realizacji projektu z EFRR w ramach Poddziałania 5.2.2.

Wsparcie inwestycji przedsiębiorstw. W zamieszczonej niżej tabeli 4 przedstawiony został obowiązujący w przypadku przeprowadzanego w latach 2008-2009 konkursu dla tego Poddziałania model preferencji ustalony przez Komitet Monitorujący, natomiast w tabeli 5 zawarte zostały średnie arytmetyczne z 3 ocen wprowadzonych do programu dla każdego z 9 kryteriów w przypadku każdego z 10 wniosków.

Tabela 4

Wagi i wartości minimalne
oraz progi równoważności, preferencji i veta dla kryteriów

Kryterium	Waga	Minimum	Próg		
			równoważności	preferencji	veta
B.2.1	0,23	2,00	1,00	3,50	7
B.2.2	0,14	2,00	0,00	3,00	11
B.2.3	0,23	0,00	0,50	2,00	8
B.2.4	0,23	0,00	1,00	3,00	8
B.2.5	0,02	0,00	1,00	3,00	11
B.2.6	0,04	2,50	1,50	3,00	11
B.2.7	0,05	2,00	1,50	3,50	11
B.2.8	0,01	0,00	3,00	5,00	11
B.3.1	0,05	0,00	0,00	2,00	7

Tabela 5

Uśrednione oceny ekspertów

Wniosek	Kryteria oceny								
	B 2.1	B 2.2	B 2.3	B 2.4	B 2.5	B 2.6	B 2.7	B 2.8	B 3.1
1	10	2	10	10	10	8,3	8,3	1,7	0,7
2	7,7	5,3	7	7,3	10	5,3	7	10	10
3	10	6,7	10	10	8,7	7	7	8,3	8,3
4	7	5,3	3,7	5	10	5	7	10	10
5	7,3	6,7	7,7	8	8,3	8	8	8,7	6
6	8	8	7,7	8,3	8,7	8,7	8,7	9	9,7
7	9	9	9,3	8	9,7	9	10	9	9,3
8	9,7	7	9,3	9,7	5,7	7,7	6,7	6,7	7,7
9	7,7	5,7	7,7	7	10	5,3	4,3	10	10
10	9	7	6,7	8	5,3	5,7	7,7	6,7	7,7

Wśród zamieszczonych w tabeli 5 wniosków o dofinansowanie realizacji projektu z UE zwraca uwagę wniosek nr 1, który otrzymał najwyższe możliwe oceny w przypadku trzech najbardziej istotnych kryteriów (B.2.1, B.2.3 i B.2.4), minimalną akceptowalną liczbę punktów w przypadku dość ważnego kryterium B.2.2 dotyczącego zgodności rozpatrywanego projektu z preferowanymi typami projektów oraz bardzo niskie oceny w przypadku kryteriów B.2.8 oraz B.3.1, z których to drugie jest o tyle ważne, że dotyczy wpływu przewidywanych rezultatów projektu na osiągnięcie założonych w RPO WK-P 2007-2013 celów.

W tabeli 6 zaprezentowane zostały wyniki uzyskane za pomocą metody PROMETHEE II z progiem veta oraz za pomocą średniej ważonej, stanowiącej metodę powszechnie stosowaną. W przypadku drugiej z wymienionych metod wniosek nr 1 sklasyfikowany został na miejscu 4, natomiast w przypadku pierwszej – dzięki zastosowaniu progów równoważności, preferencji i veta – dopiero na miejscu 9, które zdecydowanie lepiej odzwierciedla jego jakość; szczególnie wyraźnie jest to widoczne, gdy porównamy oceny przyznane rozpatrywanemu wnioskowi z ocenami wniosków nr 6 i nr 10, które to wnioski zajmują w rankingu otrzymanym za pomocą średniej ważonej odpowiednio miejsca 5 i 6. W przypadku pozostałych wniosków, mniej specyficznych niż wniosek nr 1, otrzymane uporządkowania nie różnią się od siebie.

Tabela 6

Porównanie wyników porządkowania wniosków

Numer wniosku	PROMETHEE II		Średnia ważona	
	ocena	pozycja w rankingu	ocena	pozycja w rankingu
1	-1,299	9	8,179	4
2	-1,237	8	7,164	9
3	3,338	1	9,14	1
4	-4,57	10	5,703	10
5	-1,054	6	7,501	7
6	0,932	4	8,172	5
7	3,059	2	8,918	2
8	2,631	3	8,79	3
9	-1,191	7	7,177	8
10	-0,609	5	7,602	6

Podsumowanie

W opracowaniu przedstawiono aplikację komputerową, która wykonana została na potrzeby systemu oceny projektów funkcjonującego w ramach RPO WK-P 2007-2013. Aplikacja posiada wszystkie wymagane przez zamawiającego funkcje pozwalające między innymi na wprowadzanie, przeglądanie i modyfikowanie danych dotyczących ocenianych projektów europejskich, tworzenie list wniosków ocenionych pozytywnie, negatywnie, skierowanych do ponownej oceny formalnej lub do wyjaśnienia, drukowanie protokołów oceny, a także sporządzanie rankingów wniosków przedkładanych Zarządowi Województwa.

W procedurze rankingowania projektów wykorzystana została metoda PROMETHEE II z progiem veta. Jej założenia uwzględniające m.in. niepewność i nieprecyzyjność towarzyszącą wszystkim problemom decyzyjnym zdecydowanie bardziej przystają do rzeczywistości niż założenia metody polegającej na obliczaniu średniej ważonej. Co więcej, jej zastosowanie może udoskonalić i zobiektywizować proces oceny i wyboru projektów, ograniczając – dzięki wcześniejszemu zdefiniowaniu modelu preferencji – ryzyko manipulowania wynikiem oraz wykluczając, przynajmniej częściowo, możliwość kompensacji złej oceny w przypadku jednego z kryteriów dobrą oceną w przypadku innego [5, s. 86].

**Karta oceny merytorycznej wniosku o dofinansowanie projektu współfinansowanego ze środków
Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013
w ramach 5 Osi Wzmocnienie konkurencyjności przedsiębiorstw
Poddziałanie 5.2.2. Wsparcie inwestycji przedsiębiorstw**

B2 Kryteria oceny jakości projektu					
Nr	Kryterium / Opis kryterium	Sposób określenia liczby punktów za opcje w ramach kryterium	Sposób obliczenia łącznej liczby punktów za kryterium	Minimalna liczba punktów	Punktacja
B 2.1	5. Oczekiwane rezultaty projektu w sposób bezpośredni wspierają realizację celów określonych w Uszczegółowieniu Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego dla Działania 5.2. poprzez poprawę oferty produktowej i technologicznej przedsiębiorstwa. Czy efektem projektu będzie poszerzenie oferty produktowej/usługowej lub wprowadzenie nowej (dla przedsiębiorstwa) technologii:				
	a) Wprowadzenie do oferty nowego produktu/usługi;	0 – 3 według oceny	Suma	2	
	b) Wprowadzenie do oferty udoskonalonego produktu/usługi;	0 – 3 według oceny			
	c) Wprowadzenie nowej technologii do procesu produkcji lub do procesu świadczenia usług.	0 – 4 według oceny			
B 2.2	6. Projekt jest zgodny z preferowanymi typami projektów. Projekt inwestycyjny zwiększający potencjał (produkcyjny, wytwórczy, usługowy) przedsiębiorstw przez wdrożenie rozwiązań innowacyjnych przy jednoczesnym:				
	a) Zwiększeniu zatrudnienia o 10% (minimum 1 osoba);	0 - Nie, 4 - Tak	Wartość	2	
	b) Zwiększeniu zatrudnienia o 10% - 30%;	0 - Nie, 6 - Tak			
	c) Zwiększeniu zatrudnienia powyżej 30%.	0 - Nie, 10 - Tak			
B 2.3	7. Oczekiwane rezultaty projektu w sposób bezpośredni wspierają realizację celów określonych w Uszczegółowieniu Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego dla Działania 5.2. poprzez zwiększenie efektywności działania przedsiębiorstwa. Czy projekt wpłynie na zwiększenie sprzedaży produktów/usług:				
	a) - < = 10%;	0 pkt.	Wartość	n/d	
	b) - > 10% - 20%;	1 – 3 według oceny			
	c) - > 20% - 30%;	4 – 6 według oceny			
	d) - > = 30%.	7 – 10 według oceny			

B2 Kryteria oceny jakości projektu					
Nr	Kryterium / Opis kryterium	Sposób określenia liczby punktów za opcje w ramach kryterium	Sposób obliczenia łącznej liczby punktów za kryterium	Minimalna liczba punktów	Punktacja
B 2.4	8. Oczekiwane rezultaty projektu w sposób bezpośredni wspierają realizację celów określonych w Uszczegółowieniu Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego dla Działania 5.2. poprzez zwiększenie konkurencyjności działania przedsiębiorstwa. Czy projekt wpłynie na zwiększenie sprzedaży na rynku międzynarodowym w okresie trwałości:	a) 0;	0 pkt.	Wartość	n/d
		b) od 1% do 10%;	1 – 3 według oceny		
		c) od ponad 10% do 20%;	4 – 6 według oceny		
		d) powyżej 20%.	7 – 10 według oceny		
B 2.5	9. Udział środków własnych beneficjenta w projekcie.	Czy udział środków własnych beneficjenta wynosi:			
		a) >= 50% - 60%;	1 – 5 według oceny	Wartość	n/d
		b) powyżej 60%.	6 – 10 według oceny		
B 2.6	10. Wypełnienie polityk i zasad wspólnotowych. W jakim stopniu wypełniane są właściwe polityki wspólnotowe:	a) Wyrównywanie szans kobiet i mężczyzn;	0 - 2,5 według oceny	Suma	2,5
		b) Projekt przyczynia się do społeczeństwa informacyjnego;	0 - 2,5 według oceny		
		c) Polepszenie BHP w przedsiębiorstwie;	0 - 2,5 według oceny		
		d) Projekt przyczyni się do ochrony środowiska.	0 - 2,5 według oceny		
B 2.7	11. Strategiczny charakter projektu i komplementarność z innymi przedsięwzięciami. Ocenie podlega wynikający z dokumentów strategicznych, planów wieloletnich bądź innych dokumentów planistycznych wnioskodawcy stopień komplementarności i powiązania projektu z:	a) Działaniami w ramach funduszy europejskich, w szczególności w ramach Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego 2007-2013: Oś priorytetowa 4. Działanie 4.3. Rozwój komercyjnych e-usług. Oś priorytetowa 5. Działanie 5.3. Wspieranie przedsiębiorstw w zakresie dostosowania do wymogów ochrony środowiska. Działanie 5.6.			
		0 - 2 według oceny	Suma	2	

B2 Kryteria oceny jakości projektu					
Nr	Kryterium / Opis kryterium	Sposób określenia liczby punktów za opcje w ramach kryterium	Sposób obliczenia łącznej liczby punktów za kryterium	Minimalna liczba punktów	Punktacja
	Kompleksowe uzbrojenie terenów pod inwestycje, Oś priorytetowa 6. Działanie 6.2. Rozwój usług turystycznych i uzdrowiskowych, Oś priorytetowa 7.; Programu Rozwoju Obszarów Wiejskich 2007-2013; Oś Priorytetowa 3. Działanie 3.4. Tworzenie i rozwój mikroprzedsiębiorstw; Programu Operacyjnego Innowacyjna Gospodarka 2007-2013; Oś Priorytetowa 4: Inwestycje w innowacyjne przedsięwzięcia; Programu Operacyjnego Kapitał Ludzki 2007–2013 (komponent regionalny): Wsparcie oraz promocja przedsiębiorczości i samozatrudnienia (w tym doradztwo i dotacje na rozpoczęcie własnej działalności gospodarczej);				
	b) Innymi działaniami zrealizowanymi przez wnioskodawcę ze środków własnych;	0 - 2 według oceny			
	c) Działaniami zrealizowanymi w ramach programów finansowanych ze środków Funduszy Strukturalnych w latach 2004-2006 oraz ze środków przedakcesyjnych i innych środków pomocowych;	0 - 2 według oceny			
	d) Zgodność projektu ze Strategią Rozwoju Województwa Kujawsko-Pomorskiego na lata 2007-2020 oraz Regionalną Strategią Innowacji Województwa Kujawsko-Pomorskiego do 2015 roku.	0 - 4 według oceny			
B 2.8	12. Promocja projektu.				
	Proponowane formy promocji projektu.	0 - 10 według oceny	Wartość	n/d	

B3 Kryteria oceny wpływu przewidywanych rezultatów projektu na osiągnięcie założonych w RPO celów				
Nr	Kryterium / Opis kryterium	Sposób określenia liczby punktów za opcje w ramach kryterium	Sposób obliczenia łącznej liczby punktów za kryterium	Punktacja
B 3.1	13. Wpływ projektu na rynek pracy.			
	Przewidywana liczba nowych miejsc pracy utworzonych (i utrzymywanych min. przez dwa lata) we wspieranym przedsiębiorstwie:			
	a) od 0 – 3;	0 – 3 według oceny	Wartość	
	b) od 4 – 6;	4 – 6 według oceny		
	c) powyżej 6.	7 – 10 według oceny		

Źródło: Urząd Marszałkowski Województwa Kujawsko-Pomorskiego.

Literatura

1. Brans J.P., Mareschal B., Vincke Ph. (1984). PROMETHEE: A New Family of Outranking Methods in Multicriteria Analysis. W: Operational Research '84. Red. J.P. Brans. North-Holland, Amsterdam.
2. Brans J.P., Vincke Ph. (1985). A Preference Ranking Organization Method: The PROMETHEE Method for Multiple Criteria Decision-Making. Management Science, 31.
3. Brans J.P., Vincke Ph., Mareschal B. (1986). How to Select and How to Rank Projects The PROMETHEE Method. European Journal of Operational Research, 24.
4. Górecka D. (2009). Wielokryterialne wspomaganie wyboru projektów europejskich. TNOiK „Dom Organizatora”, Toruń.
5. Górecka D. (2010). Wykorzystanie metod wielokryterialnych w procesie oceny i wyboru wniosków o dofinansowanie realizacji projektu z funduszy Unii Europejskiej. W: Współczesne tendencje rozwojowe badań operacyjnych. Red. J. Siedlecki, P. Peternek. Uniwersytet Ekonomiczny, Wrocław.
6. Górecka D. (2010). Zastosowanie metod wielokryterialnych opartych na relacji przewyższania do oceny europejskich projektów inwestycyjnych. W: Metody i zastosowania badań operacyjnych'10. Red. M. Nowak. Uniwersytet Ekonomiczny, Katowice.
7. Roy B. (1990). Wielokryterialne wspomaganie decyzji. Wydawnictwa Naukowo-Techniczne, Warszawa.
8. System oceny projektów w ramach Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013. Załącznik do uchwały nr 56/911/09 Zarządu Województwa Kujawsko-Pomorskiego z dnia 21 lipca 2009 roku. <http://www.mojregion.eu/regionalny-program-operacyjny-wojewodztwa-kujawsko-pomorskiego/wazne-dokumenty/wytyczne-instytucji-zaradzajacej/inne.html>.
9. Szczegółowy opis osi priorytetowych Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013 (Uszczegółowienie RPO) – wersja 4.8 z 24 lutego 2011 roku. <http://www.mojregion.eu/regionalny-program-operacyjny-wojewodztwa-kujawsko-pomorskiego/wazne-dokumenty/dokumenty-programowe/szczegolowy-opis-osi-prioryt.html>.
10. Załącznik V Kryteria wyboru projektów do Szczegółowego opisu osi priorytetowych Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego (URPO WK-P) na lata 2007-2013 z dnia 24 lipca 2008. [//fundusze.kujawsko-pomorskie.pl](http://fundusze.kujawsko-pomorskie.pl).

**APPLICATION OF PROMETHEE II IN THE EUROPEAN PROJECTS' RANKING
PROCESS WITHIN THE REGIONAL OPERATIONAL PROGRAMME
FOR KUJAWSKO-POMORSKIE VOIVODESHIP FOR THE YEARS 2007-2013**

Summary

In the article a computer application created for Marshal Office of Kujawsko-Pomorskie Voivodship, enabling the use of PROMETHEE II method with a veto threshold in the process of ranking projects applying for co-financing from the European Union funds in the framework of the Regional Operational Programme for Kujawsko-Pomorskie Voivodeship for the years 2007-2013 (RPO WK-P 2007-2013) will be presented.

In accordance with the expectations of the decision-maker the programme interface allows entering data relating to applications for project co-financing submitted in response to advertised under RPO WK-P 2007-2013 competitions, setting their status as well as creating lists of applications evaluated positively, negatively and submitted to a formal re-evaluation or to clarification. The most important function of the programme is ordering projects with the help of PROMETHEE II method with a veto threshold and building ranking lists, in accordance with which the EU financial assistance is granted.

Piotr Jaśkowski

Robert Bucoń

Sławomir Biruk

Politechnika Lubelska

OCENA PRZYDATNOŚCI UŻYTKOWEJ BUDYNKÓW MIESZKALNYCH Z ZASTOSOWANIEM ROZMYTEGO ROZSZERZENIA AHP*

Wprowadzenie

Dokonywanie ocen stanu przydatności użytkowej budynków mieszkalnych i na tej podstawie formułowanie wytycznych w zakresie programu działań modernizacyjnych i remontowych ma istotne znaczenie w procesie zarządzania nieruchomościami. Ocena przydatności użytkowej budynków mieszkalnych wymaga określenia wymagań eksploatacyjnych (kryteriów oceny) i ich wpływu na przydatność użytkową oraz dokonania ocen częściowych (określenia stopnia spełnienia wymagań). Z przeprowadzonych wywiadów z użytkownikami, zarządcami, jak również ekspertami zajmującymi się oceną stanu budynków mieszkalnych wynika, że wymagania te odnoszą się nie tylko do cech związanych ze stanem technicznym, ale również są istotne właściwości energetyczne, funkcjonalne oraz stan wizualny.

Na ocenę przydatności użytkowej budynku mają wpływ wagi kryteriów przyjętych do analizy. Wyrażają one w sposób wymierny preferencje uczestników procesu eksploatacyjnego. Ze względu na konieczność posiadania dużego doświadczenia i wiedzy specjalistycznej z różnych dziedzin, wskazanie konkretnych wartości wag kryteriów jest trudne i wymaga często uwzględnienia opinii wielu osób (użytkownicy, eksperci, projektanci, zarządcy, rzeczoznawcy).

Autorzy proponują zastosowanie w procesie grupowego podejmowania decyzji rozmytego rozszerzenia metody AHP, które umożliwi ustalenie istotności kryteriów poprzez agregację ocen grupy ekspertów.

* Praca została sfinansowana przez Ministerstwo Nauki i Szkolnictwa Wyższego (grant nr N N506 254637).

1. Metoda AHP i jej rozmyte modyfikacje

Jedną z metod wspomagających ustalanie istotności kryteriów (i również ocenę wariantów) jest metoda AHP (analityczny proces hierarchiczny). Procedura analizy obejmuje następujące etapy: tworzenie hierarchii celów problemu decyzyjnego, porównywanie parami istotności kryteriów na każdym poziomie hierarchii (przy zastosowaniu skali 9-stopniowej), ustalenie wag kryterium (wektora własnego macierzy porównań odpowiadającego maksymalnej wartości własnej).

Podobnie, gdy kryteria mają charakter jakościowy, przeprowadza się porównywanie parami wariantów dla poszczególnych kryteriów, co umożliwia w wyniku obliczenia ocen cząstkowych i oceny syntetycznej (jako sumy ważonej ocen cząstkowych).

Metoda AHP, opracowana przez Thomasa Saaty'ego w latach 80. XX wieku [14; 15], jest prezentowana w literaturze jako efektywne narzędzie wspomagania decyzji przy wielorakości atrybutów oceny w wielu obszarach zarządzania przedsięwzięciem budowlanym (między innymi przy selekcji ofert na roboty budowlane [10], wyborze technologii procesów [13], doborze maszyn [7], ocenie procedury określania wielkości budżetu przedsięwzięcia [9]).

Od ponad dwudziestu lat wielu autorów kwestionuje założenia metody AHP, wskazując jej ograniczenia i wady (między innymi dotyczące stosowanej skali ocen, zasadności obliczania maksymalnej wartości własnej i odpowiadającego mu wektora własnego macierzy porównań przy braku zgodności opinii). Znanym problemem prezentowanym w literaturze jest zmiana rankingu wariantów w przypadku dodania nowych alternatyw [1].

Metoda ta nie umożliwia w sposób bezpośredni uwzględnienia warunków ryzyka związanego z niepełnym dostępem do danych, informacji, subiektywizmem ocen. Decydentowi w wielu przypadkach łatwiej jest określić względne preferencje poprzez wskazanie zakresu wartości lub opisowo – niż podając konkretne wartości z przyjętej skali ocen. Ostre wartości preferencji (z 9-stopniowej skali) nie umożliwiają modelowania niepewności oszacowań dokonywanych przez decydenta, dlatego są podejmowane prace nad modyfikacją klasycznej metody AHP, poprzez np. umożliwienie wyrażania opinii w postaci liczb rozmytych.

Pierwsze rozmyte rozwinięcia metody AHP opracowali Van Laarhoven i Pedrycz [17] oraz Buckley [2] i Chang [4], umożliwiając ocenę względnych preferencji za pomocą liczb rozmytych o trójkątnych funkcjach przynależności $\tilde{a}_{ij} = (l_{ij}, m_{ij}, u_{ij})$. Decydent dokonując porównania parami ważności kryteriów wyraża swoje preferencje podając wartość „około m_{ij} ”, przy czym swoją niepewność co do precyzji tego określenia wyraża w postaci długości przedziału

$[l_{ij}, u_{ij}]$. Porównywanie parami jest możliwe również przy zastosowaniu fuzyfikacji określeń lingwistycznych (skal przypisujących im liczby rozmyte o dowolnych funkcjach przynależności).

Złożonym obliczeniowo problemem w rozmytych rozwinięciach metody jest procedura obliczania wektora wag na podstawie macierzy porównań istotności parami. Stosowane są przy tym między innymi metoda średniej geometrycznej [2], metoda rozmytych najmniejszych kwadratów [17], metoda lambda-max (stanowiąca rozmyte rozwinięcie koncepcji wektora własnego) [6] oraz rozmyte programowanie preferencji [11], które autorzy proponują zastosować w procedurze obliczania wag kryteriów na podstawie macierzy zagregowanych ocen ekspertów.

Mikhailov [11] przeanalizował trudności związane z konstruowaniem macierzy porównań z rozmytymi wartościami odwrotnymi i możliwość uzyskania błędnych wyników. Według Saaty'ego [15] oceny pośrednie są wystarczającym środkiem wyrażenia niepewnej opinii w przypadku wątpliwości decydenta, a rozmyte oceny jedynie zwiększają zgodność porównań kosztem mniejszej dokładności oszacowań wag i w efekcie mogą doprowadzić do błędnych decyzji. W przypadku wyrażania rozmytych preferencji, istnieje nieskończona liczba wartości własnych i wektorów własnych macierzy porównań, a problem wyboru najlepszego rozwiązania musi być rozpatrywany jako poszukiwanie kompromisu między wartością wskaźnika niezgodności a wartością funkcji przystosowania. Duży poziom zgodności ocen nie gwarantuje dużego poziomu wiarygodności wyników [8].

2. Agregacja ocen wielu decydentów

Celem agregacji opinii ekspertów jest wyznaczenie jednej wspólnej opinii kolektywnej, będącej pewnym kompromisem, ze względu na brak zgodności preferencji decydentów.

Rozważmy problem określenia ważności n kryteriów (lub analogicznie n obiektów budowlanych). W procesie decyzyjnym uczestniczy K ekspertów. Każdy z nich dokonuje $m = n(n-1)/2$ porównań parami kryteriów na danym poziomie hierarchii problemu (określa względny stopień przewyższania, preferencji, ważności), stosując skalę $1/9, 1/7, 1/5, 1/3, 1, 3, 5, 7, 9$ rozszerzoną ewentualnie o oceny pośrednie $1/8, 1/6, 1/4, 1/2, 2, 4, 6, 8$.

W wyniku operacji porównywania parami powstaje zbiór K macierzy $A_k = \{a_{ijk}\}$, $i = 1, 2, \dots, n-1$, $j = 2, 3, \dots, n$, $j > i$, $k = 1, 2, \dots, K$, gdzie a_{ijk} oznacza ocenę preferencji kryterium i względem j (czyli iloraz wag kryteriów i i j), wyrażoną w przyjętej skali ocen i dokonaną przez eksperta k .

W klasycznej metodzie AHP agregacji opinii dokonuje się poprzez ich uśrednianie geometryczne. Średnia arytmetyczna nie może być stosowana do agregacji opinii ekspertów [5]. Stosowanie średniej geometrycznej umożliwia zapewnienie spełnienia warunku optymalności Pareto (alternatywa preferowana przez każdego eksperta powinna być preferowana przez całą grupę), lecz tylko przy małej rozbieżności opinii [16]. Średnia geometryczna wyraża pewien konsensus grupy, ale nie w pełni odzwierciedla rozbieżność preferencji ekspertów.

Chang i inni [3] zaproponowali wyrażać oceny zagregowane w postaci liczb rozmytych $\tilde{a}_{ijk} = (l_{ij}, m_{ij}, u_{ij})$ o trójkątnych funkcjach przynależności $\mu_{a_{ij}}(x) \in [0,1]$, których punkty charakterystyczne wyznacza się z zależności

$$l_{ij} = \min_{k=1,2,\dots,K} \{a_{ijk}\} \quad (1)$$

$$m_{ij} = \left(\prod_{k=1}^K a_{ijk} \right)^{1/K} \quad (2)$$

$$u_{ij} = \max_{k=1,2,\dots,K} \{a_{ijk}\} \quad (3)$$

Takie podejście w konstruowaniu funkcji przynależności jest oparte na założeniu o równomiernym rozkładzie preferencji ekspertów, które w praktyce rzadko jest spełnione.

Mikhailov [12] opracował rozmyty model programowania liniowego z funkcją celu, za pomocą której jest maksymalizowana zgodność spełnienia dla każdego eksperta równania $w_i/w_j = a_{ijk}$, z uwzględnieniem dopuszczalnej tolerancji d_{ijk} (dopuszczalna wielkość różnicy pomiędzy ilorazem w_i/w_j a wyrażoną opinią a_{ijk}). Ponieważ wielkości parametrów d_{ijk} , dla każdego eksperta i każdej opinii, mogą znacząco wpływać na wynik końcowy, konieczne jest opracowanie wiarygodnych zasad ich określania [8].

3. Proponowana procedura obliczania wag kryteriów

Celem zastosowania tej procedury jest określenie ostrych wartości wag kryteriów $w = [w_1, w_2, \dots, w_n]^T$ na podstawie wcześniej dokonanych porównań i ich agregacji.

Proponuje się modelować zagregowane opinie za pomocą liczb rozmytych $\tilde{a}_{ij}$ z funkcjami przynależności $\mu_{a_{ij}}(x) \in [0,1]$, tworzonymi w sposób następujący:

1. Przedział $[l_{ij}, m_{ij}]$ należy podzielić na t przedziałów o jednakowej długości: $[s_0^-, s_1^-], [s_1^-, s_2^-], \dots, [s_{t-1}^-, s_t^-]$, $s_t^- = m_{ij}$, $s_0^- = l_{ij}$
2. Przedział $[m_{ij}, u_{ij}]$ należy podzielić na t przedziałów o jednakowej długości: $[s_1^+, s_2^+], [s_2^+, s_3^+], \dots, [s_t^+, u_{ij}]$, $s_1^+ = m_{ij}$, $u_{ij} = s_{t+1}^+$
3. Wartości funkcji przynależności należy obliczyć z następujących zależności

$$\mu_{a_{ij}}^-(s_l^-) = \frac{g_l^-}{v^-}, \quad l = 1, 2, \dots, t \quad (4)$$

$$\mu_{a_{ij}}^+(s_l^+) = \frac{g_l^+}{v^+}, \quad l = 1, 2, \dots, t \quad (5)$$

$$\mu_{a_{ij}}^-(x) = \mu_{a_{ij}}^-(s_{l-1}^-) - \left[\mu_{a_{ij}}^-(s_l^-) - \mu_{a_{ij}}^-(s_{l-1}^-) \right] \frac{x - s_{l-1}^-}{s_l^- - s_{l-1}^-}, \quad \forall x \in (s_{l-1}^-, s_l^-) \quad (6)$$

$$\mu_{a_{ij}}^+(x) = \mu_{a_{ij}}^+(s_l^+) - \left[\mu_{a_{ij}}^+(s_l^+) - \mu_{a_{ij}}^+(s_{l+1}^+) \right] \frac{x - s_l^+}{s_{l+1}^+ - s_l^+}, \quad \forall x \in (s_l^+, s_{l+1}^+) \quad (7)$$

$$\mu_{a_{ij}}^-(x) = 0, \quad \forall x \notin (l_{ij}, u_{ij}) \quad (8)$$

gdzie:

v^- – liczba opinii a_{ijk} mniejszych równych od m_{ij} ,

v^+ – liczba opinii a_{ijk} większych równych od m_{ij} ,

g_l^- – liczba opinii a_{ijk} należących do przedziału $[l_{ij}, s_l^-]$,

g_l^+ – liczba opinii a_{ijk} należących do przedziału $[s_l^+, u_{ij}]$.

Procedura ta jest oparta na metodyce reprezentacji liczb rozmytych za pomocą α -przekrojów trójkątnych funkcji przynależności

$$S_{ij}^\alpha = \left\{ x \in \mathbb{R} : \mu_{a_{ij}}^-(x) \geq \alpha \right\} = [l_{ij}(\alpha), u_{ij}(\alpha)], \quad \forall \alpha \in (0, 1] \quad (9)$$

$$S_{ij}^0 = [l_{ij}, u_{ij}] \quad (10)$$

Wektor wag kryteriów dla każdego przekroju powinien być ustalony w sposób taki, aby spełniona była nierówność (w sensie rozmytym)

$$l_{ij}(\alpha) \leq \frac{w_i}{w_j} \leq u_{ij}(\alpha), \quad i = 1, 2, \dots, n-1, j = 2, 3, \dots, n, j > i \quad (11)$$

lub w postaci równoważnej następujące warunki

$$\begin{cases} w_i - w_j u_{ij}(\alpha) \leq 0 \\ -w_i + w_j l_{ij}(\alpha) \leq 0 \end{cases}, \quad i = 1, 2, \dots, n-1, j = 2, 3, \dots, n, j > i \quad (12)$$

Powyższy układ $2m$ nierówności można zapisać w układzie macierzowym w postaci $Rw \leq 0$, gdzie macierz $R \in \mathfrak{R}^{2m \times n}$. Nierówność s ($R_s w \leq 0$, $s = 1, 2, \dots, 2m$) jest rozmytym liniowym ograniczeniem, którego spełnienie można opisać za pomocą następującej funkcji przynależności [11]

$$\mu_s(R_s w) = \begin{cases} 1 - \frac{R_s w}{d}, & R_s w < d \\ 0, & R_s w \geq d \end{cases} \quad (13)$$

gdzie d – jest tolerancją.

Funkcja przynależności przyjmuje wartość 0, gdy nierówność s jest niespełniona, wartości z przedziału $(0, 1)$, gdy $0 < R_s w < d$, wartości z przedziału $[1, +\infty]$, gdy $R_s w \leq 0$. Przynależność rozwiązania (wektora wszystkich wag) do obszaru rozwiązań dopuszczalnych $\tilde{P}$ (w sensie rozmytym) określa funkcja

$$\mu_p(w) = [\min\{\mu_1(R_1 w), \dots, \mu_m(R_m w)\} \mid w_1 + \dots + w_n = 1] \quad (14)$$

Według Mikhailova [11] rozwiązaniu najbardziej satysfakcjonującemu powinna odpowiadać maksymalna wartość funkcji przynależności do zbioru rozwiązań dopuszczalnych. Autorzy proponują jednak zastosowanie innego, ale równoważnego podejścia, polegającego na minimalizacji maksymalnych wartości funkcji przynależności $\mu_s(R_s w)$, co zapewnia maksymalizację minimalnych wartości funkcji $\mu_{s+1}(R_{s+1} w)$ dla s nieparzystych i maksymalizację minimalnych wartości funkcji $\mu_{s-1}(R_{s-1} w)$ dla s parzystych. Problem ten można sformułować w postaci zadania programowania liniowego [11]

$$\begin{cases} \min z = \lambda \\ d\lambda + R_s w \geq d, \quad s = 1, 2, \dots, 2m \\ \sum_{i=1}^n w_i = 1 \\ w_i > 0, \quad i = 1, 2, \dots, n \\ \lambda \geq 0. \end{cases} \quad (15)$$

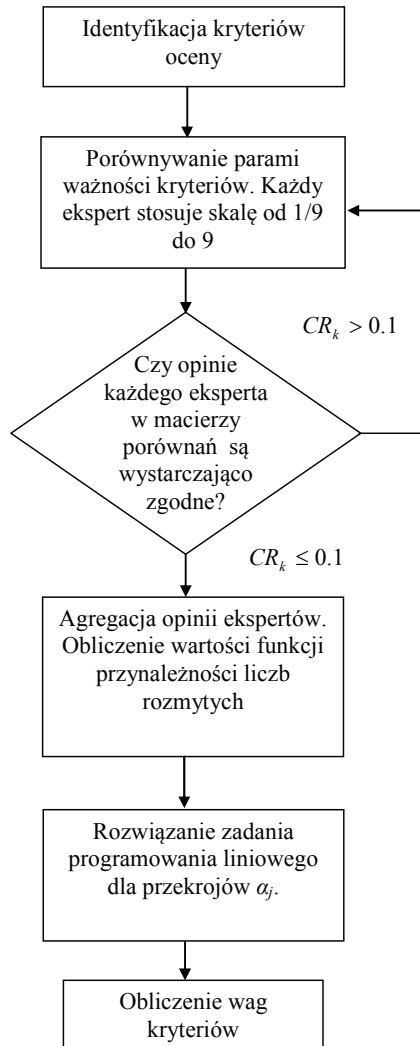
Takie sformułowanie problemu jest równoważne z modelem zaproponowanym przez Mikhailova [12] do agregacji opinii ekspertów, w którym zakłada się jednakowe wartości wszystkich parametrów d_{ijk} .

Powyższe zadanie można rozwiązać dla skończonej liczby L przekrojów, uzyskując w wyniku optymalne wartości wektora wag $w^*(\alpha_j)$ i zmiennej $\lambda^*(\alpha_j)$. Wartość α_j może być traktowana jako istotność danego rozwiązania (małym wartościom odpowiada szeroki zakres opinii; dla dużych wartości uzyskane rozwiązanie powinno wyrażać konsensus grupy ekspertów), zatem zagregowaną wartość wag kryteriów można obliczyć ze wzoru [11]

$$\bar{w}_i = \sum_{j=1}^L \alpha_j w_i^*(\alpha_j) / \sum_{j=1}^L \alpha_j, \quad i = 1, 2, \dots, n \quad (16)$$

Wartość zmiennej λ określa stopień satysfakcji z uzyskanego rozwiązania (w sensie spełnienia wszystkich warunków), a także stopień zgodności względnych preferencji. Oceny można uznać za zgodne, a rozwiązanie za w pełni satysfakcjonujące, gdy $\lambda \geq 1$. W przypadku, gdy $0 < \lambda < 1$, rozwiązanie spełnia wszystkie ograniczenia z uwzględnieniem tolerancji d . Dla małych wartości d zadanie może być sprzeczne. Zbiór rozwiązań dopuszczalnych nigdy nie jest pusty dla $d \geq 1$, a wielkość tolerancji nie wpływa na wartość rozwiązania optymalnego [8, 11]. Jakość otrzymywanych wyników, w porównaniu z wynikami uzyskiwanymi przy zastosowaniu innych metod agregacji opinii ekspertów, analizowano w artykule [8].

Na rys. 1 przedstawiono kolejne etapy proponowanej procedury ustalania wag kryteriów.



Rys. 1. Proponowana procedura ustalania wag kryteriów [8]

4. Przykład

Przedstawioną metodę zastosowano do ustalenia ważności kryteriów oceny przydatności użytkowej budynków mieszkalnych. Zidentyfikowano następujące kryteria oceny:

1. Funkcjonalność.
2. Energooszczędność.
3. Stan techniczny.
4. Wygląd.

W celu ustalenia ważności tych kryteriów o opinie poproszono 15 ekspertów. Wartości względnych preferencji poszczególnych ekspertów zestawiono w tabeli 1. Na tej podstawie obliczono, zgodnie z zależnościami (4)-(8), wartości funkcji przynależności liczb rozmytych $\tilde{a}_{ij}$. Wartości funkcji przynależności dla $s_l^-, s_l^+, l=1, 2, 3$ zestawiono w tabeli 2. Na rys. 2 przedstawiono wykres funkcji przynależności liczby $\tilde{a}_{13}$.

Tabela 1

Względne preferencje ekspertów dla poszczególnych par kryteriów

Numer kryterium		Numer eksperta k														
i	j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2	1/4	1/2	1/3	1/4	1/4	1/2	1	1/2	1/3	1/3	1/2	1/2	1/2	1/2	1/3
1	3	1/5	1/6	1/6	1/5	1/7	1/6	1/5	1/4	1/3	1/4	1/7	1/4	1/4	1/7	1/4
1	4	4	3	4	5	5	4	4	4	6	6	3	5	7	4	3
2	3	1/4	1/3	1/4	1/3	1/2	1/3	1/3	1/4	1/2	1/4	1/4	1/4	1/2	1/4	1/4
2	4	5	6	5	6	7	5	8	4	9	5	7	8	8	7	5
3	4	8	9	9	9	9	7	8	9	9	7	9	9	8	7	9

Rozwiązania optymalne zadań (15) dla różnych przekrojów α_j i wartości tolerancji d zestawiono w tabeli 3. Ponieważ wszystkie wartości $\lambda^* > 1$, to uzyskane rozwiązania spełniają wszystkie ostre ograniczenia zadania.

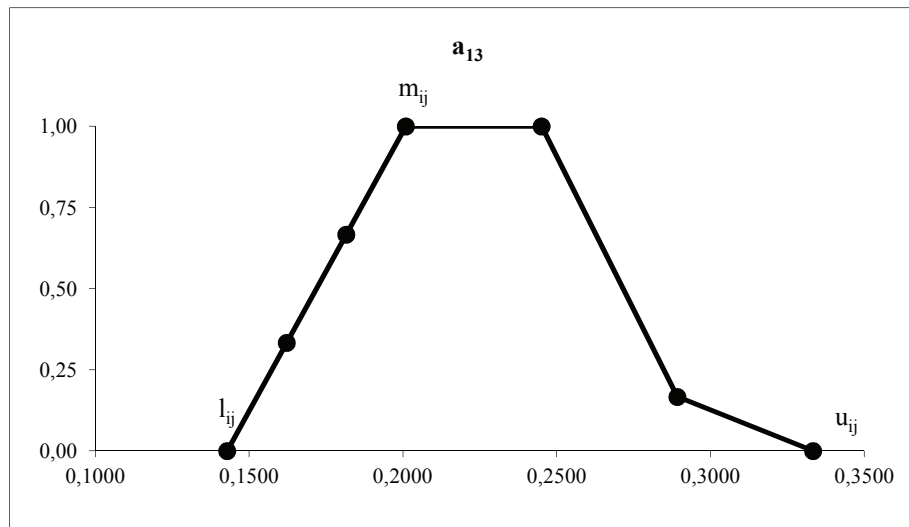
Rys. 2. Wykres funkcji przynależności dla zagregowanej opinii $\tilde{a}_{13}$

Tabela 2

Wartości funkcji przynależności dla zagregowanych opinii $\tilde{a}_{ij}$

i	j	$\frac{\mu_{\sim}(s_0^-)}{a_{ij}}$ s_0^-	$\frac{\mu_{\sim}(s_1^-)}{a_{ij}}$ s_1^-	$\frac{\mu_{\sim}(s_2^-)}{a_{ij}}$ s_2^-	$\frac{\mu_{\sim}(s_3^-)}{a_{ij}} = \frac{\mu_{\sim}(s_1^+)}{a_{ij}}$ s_3^- s_1^+	$\frac{\mu_{\sim}(s_2^+)}{a_{ij}}$ s_2^+	$\frac{\mu_{\sim}(s_3^+)}{a_{ij}}$ s_3^+	$\frac{\mu_{\sim}(s_4^+)}{a_{ij}}$ s_4^+
1	2	0.0000	0.4286	1.0000	1.0000	0.1250	0.1250	0.0000
		0.2500	0.3030	0.3561	0.4091	0.6061	0.8030	1.0000
1	3	0.0000	0.3333	0.6667	1.0000	1.000	0.1667	0.0000
		0.1429	0.1622	0.1816	0.2009	0.2451	0.2892	0.3333
1	4	0.0000	0.3333	0.6667	1.0000	0.5000	0.1667	0.0000
		3.0000	3.4422	3.8843	4.3265	5.2177	6.1088	7.0000
2	3	0.0000	1.0000	1.0000	1.0000	0.4286	0.4286	0.0000
		0.2500	0.2700	0.2900	0.3101	0.3734	0.4367	0.5000
2	4	0.0000	0.1250	0.7500	1.0000	0.5714	0.1429	0.0000
		4.0000	4.7226	5.4452	6.1679	7.1119	8.0560	9.0000
3	4	0.0000	0.5000	0.5000	1.0000	1.0000	1.0000	0.0000
		7.0000	7.4532	7.9064	8.3596	8.5730	8.7865	9.0000

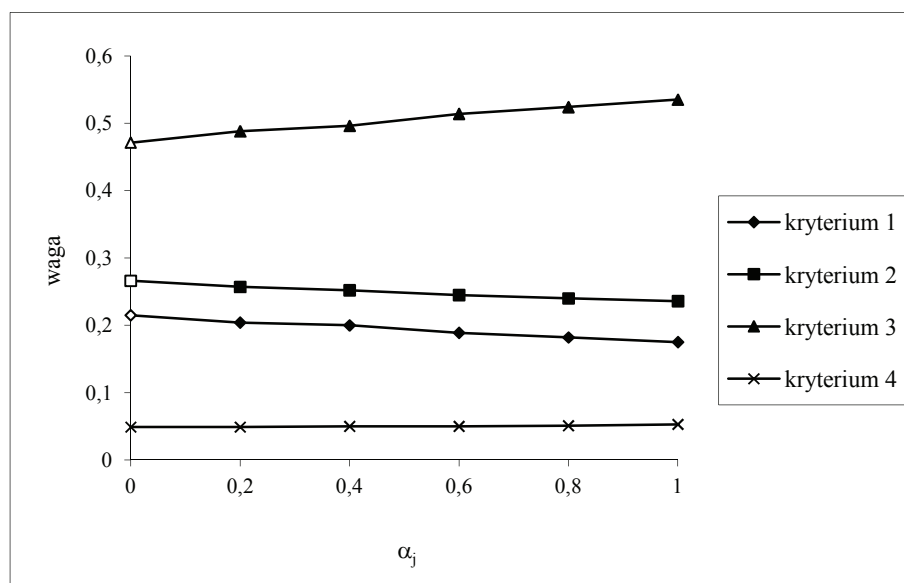
Rys. 3. Wagi kryteriów dla różnych przekrojów α_j

Tabela 3

Rozwiązania optymalne przykładu

α_j	d			
	0.5		1.0	
0	$\lambda^*=1.297$ $w_1^*=0.215$ $w_3^*=0.471$	$w_2^*=0.266$ $w_4^*=0.046$	$\lambda^*=1.148$ $w_1^*=0.215$ $w_3^*=0.471$	$w_2^*=0.266$ $w_4^*=0.046$
0.2	$\lambda^*=1.267$ $w_1^*=0.204$ $w_3^*=0.488$	$w_2^*=0.257$ $w_4^*=0.049$	$\lambda^*=1.133$ $w_1^*=0.204$ $w_3^*=0.488$	$w_2^*=0.257$ $w_4^*=0.049$
0.4	$\lambda^*=1.249$ $w_1^*=0.200$ $w_3^*=0.496$	$w_2^*=0.252$ $w_4^*=0.050$	$\lambda^*=1.124$ $w_1^*=0.200$ $w_3^*=0.496$	$w_2^*=0.252$ $w_4^*=0.050$
0.6	$\lambda^*=1.221$ $w_1^*=0.189$ $w_3^*=0.514$	$w_2^*=0.245$ $w_4^*=0.050$	$\lambda^*=1.110$ $w_1^*=0.189$ $w_3^*=0.514$	$w_2^*=0.245$ $w_4^*=0.050$
0.8	$\lambda^*=1.202$ $w_1^*=0.182$ $w_3^*=0.524$	$w_2^*=0.240$ $w_4^*=0.051$	$\lambda^*=1.101$ $w_1^*=0.182$ $w_3^*=0.524$	$w_2^*=0.240$ $w_4^*=0.051$
1.00	$\lambda^*=1.182$ $w_1^*=0.175$ $w_3^*=0.535$	$w_2^*=0.236$ $w_4^*=0.053$	$\lambda^*=1.091$ $w_1^*=0.175$ $w_3^*=0.535$	$w_2^*=0.236$ $w_4^*=0.053$

Zmiana wartości d nie powoduje zmian wartości wag kryteriów dla danego przekroju. Również iloczyn $d \cdot (1 - \lambda^*)$ w każdym przekroju przyjmuje stałą wartość. Wagi kryteriów dla różnych poziomów przekrojów przedstawiono na rys. 3. Hierarchia ważności kryteriów jest taka sama dla wszystkich przekrojów. Zagregowane wartości wag są równe: $\bar{w}_1 = 0.1849$, $\bar{w}_2 = 0.2424$, $\bar{w}_3 = 0.5195$, $\bar{w}_4 = 0.0512$.

Podsumowanie

Rozmyte rozwinięcie metody AHP pozwala na jej zastosowanie zarówno do wspomagania decyzji grupy ekspertów, jak i w przypadku wyrażania względnych preferencji w postaci liczb rozmytych i niepewności co do precyzji wyrażanych opinii. Uzyskiwane wyniki umożliwiają bardziej szczegółową analizę problemu decyzyjnego w zakresie dokładności oszacowań i zgodności ocen.

Proponowana metoda nie wymaga tworzenia pełnej macierzy porównań. W wyniku uzyskiwane są ostre wartości wag kryteriów, które można stosować przy ocenie wariantów za pomocą znanych metod MADM.

Proponowane rozmyte rozwinięcie metody AHP stanowi narzędzie wspomagające ustalenie istotności kryteriów branych pod uwagę przy określaniu stanu eksploatacyjnego obiektu mieszkalnego. Ze względu na złożoność obliczeń, prezentowany model ograniczono do analizy istotności tylko 4 kryteriów nadrzędnych. Zaproponowana metoda może być zastosowana do agregacji preferencji ekspertów dla większej liczby kryteriów na każdym poziomie hierarchii problemu decyzyjnego.

Literatura

1. Barzilai J. (1997). Deriving Weights from Pairwise Comparison Matrices. *Journal of the Operational Research Society*, 48(12), s. 1226-1232.
2. Buckley J. (1985). Fuzzy Hierarchical Analysis. *Fuzzy Sets and Systems*, 17, s. 233-247.
3. Chang Ch.-W., Wu Ch.-R., Lin H.-H. (2009). Applying Fuzzy Hierarchy Multiple Attributes to Construct an Expert Decision Making Process. *Expert Systems with Applications*, 36, s. 7363-7368.
4. Chang D.-Y. (1996). Applications of the Extent Analysis Method on Fuzzy AHP. *European Journal of Operational Research*, 95, s. 649-655.
5. Cho Y.G., Cho K.T. A Loss Function Approach to Group Preference Aggregation in the AHP. *Computers & Operations Research*, 2008, 35, s. 884-892.

6. Csutora R., Buckley J.J. (2001). Fuzzy Hierarchical Analysis: The Lambda-Max Method. *Fuzzy Sets and Systems*, 120, s. 181-195.
7. Goldenberg M., Shapira A. (2007). Systematic Evaluation of Construction Equipment Alternatives: Case Study. *Journal of Construction Engineering and Management*, 133(1), s. 72-85.
8. Jaśkowski P., Biruk S., Bucoń R. (2010). Assessing Contractor Selection Criteria Weights with Fuzzy AHP Method Application in Group Decision Environment. *Automation in Construction*, 19(2), s. 120-126 (wykorzystano za zgodą wydawnictwa Elsevier).
9. Lai Y.-T., Wang W.-Ch., Wang H.H. (2008). AHP- and Simulation-based Budget Determination Procedure for Public Building Construction Projects. *Automation in Construction*, 17, s. 623-632.
10. Lin Ch.-Ch., Wang W.-Ch., Yu W.-D. (2008). Improving AHP for Construction with an Adaptive AHP approach (A^3). *Automation in Construction*, 17, s. 180-187.
11. Mikhailov L. (2003). Deriving Priorities from Fuzzy Pairwise Comparison Judgments. *Fuzzy Sets and Systems*, 134, s. 365-385.
12. Mikhailov L. (2004). Group Prioritization in the AHP by Fuzzy Preference Programming Method. *Computers & Operations Research*, 31, s. 293-301.
13. Pan N.-F. (2008). Fuzzy AHP Approach for Selecting the Suitable Bridge Construction Method. *Automation in Construction*, 17, s. 958-965.
14. Saaty T.L., Vargas L.G. (2007). Dispersion of Group Judgments. *Mathematical and Computer Modelling*, 46, s. 918-925.
15. Saaty T.L., Tran L.T. (2007). On the Invalidity of Fuzzifying Numerical Judgments in the Analytic Hierarchy Process. *Mathematical and Computer Modelling*, 46, s. 962-975.
16. Van Den Honert R.C., Lootsma F.A. (1996). Group Preference Aggregation in the Multiplicative AHP. The Model of the Group Decision Process and Pareto Optimality. *European Journal of Operational Research*, 96, s. 363-370.
17. Van Laarhoven P.J.M., Pedrycz W. (1983). Fuzzy Extension for Saaty's Priority Theory. *Fuzzy Sets and Systems*, 11, s. 229-241.

ASSESSING RELATIVE WEIGHTS OF RESIDENTIAL BUILDING USABILITY CRITERIA BY MEANS OF FUZZY AHP

Summary

Quality of a residential building is conditioned not only by technical condition of the structure, services and finishings, but also the level of fulfilling economic, ergonomic and aesthetic requirements. These are not always of formal character, and their assessment cannot be based on dichotomous scales, so without expressing to what extent the requirements of building

users, owners, managers or valuation surveyors are met. Assessment of a building's usability involves also determining significance of the requirements. The authors argue that applying fuzzy AHP – a technique invented to deal with imprecise and uncertain information obtained from a number of sources – facilitates aggregation of opinions and finding a compromise on the building usability criteria weights.

Leszek Klukowski

Instytut Badań Systemowych PAN w Warszawie

WŁASNOŚCI ESTYMATORÓW RELACJI PORZĄDKU OPARTYCH NA RÓŻNICACH RANG – BADANIE SYMULACYJNE*

Wprowadzenie

W pracach Klukowskiego [3; 4; 5], zaproponowano estymatory relacji porządku na podstawie porównań parami w postaci różnicy rang, z błędami losowymi. Są one rozwinięciem koncepcji nearest adjoining order [1; 2; 6]. Rozważono dwie postacie tych estymatorów: pierwszy polega na minimalizacji sumy modułów różnic między postacią relacji, określoną przy użyciu pewnej funkcji, oraz porównaniami, drugi – na minimalizacji sumy modułów między postacią relacji a medianami (wartościami środkowymi) z porównań każdej pary. Własności tych estymatorów, określone w sposób analityczny, dotyczą głównie własności asymptotycznych – nie obejmują precyzji (rozkładów błędów) estymacji. Uzyskanie tych rozkładów nie wydaje się możliwe, natomiast ich znajomość jest niezbędna w zastosowaniach.

Celem niniejszej pracy jest przedstawienie wyników badań symulacyjnych w tym zakresie. Obejmują one określenie rozkładów częstości składowych błędów estymacji obu rozważanych estymatorów, w przypadku wybranych rozkładów prawdopodobieństw błędów porównań parami. Podstawę badań stanowi zbiór 9-elementowy, w którym określono trzy postacie relacji porządku oraz trzy rozkłady błędów porównań. Wykonano 200 symulacji każdego rozważanego przypadku. Wnioski z badań umożliwiają: wyznaczenie liczby porównań każdej pary elementów, która zapewnia niezbędną precyzję estymacji, a w szczególności błąd zerowy.

Praca składa się z pięciu części. W części pierwszej sformułowano problem estymacji, w drugiej określono postać estymatorów. Części – trzecia i czwarta zawierają – parametry badania symulacyjnego oraz ich wyniki. Ostatnia część pracy zawiera podsumowanie.

* Praca wykonana w ramach Projektu Badawczego Nr N N111 434937 Min. Nauki i Szkolnictwa Wyższego: Metody estymacji relacji: porządku, równoważności i tolerancji w zbiorze skończonym.

1. Sformułowanie problemu estymacji

Problem estymacji relacji porządku na podstawie porównań parami zakłóconych błędami losowymi można sformułować w następujący sposób. Dany jest zbiór $\mathbf{X} = \{x_1, \dots, x_m\}$ ($m \geq 3$), w którym istnieje relacja słabego porządku R o postaci

$$R = I \cup P \quad (1)$$

gdzie:

I – relacja równoważności (zwrotna, przechodnia, symetryczna),

P – relacja porządku (przechodnia, antysymetryczna).

Relacja R generuje uporządkowaną rodzinę podzbiorów $\mathcal{X}_1^*, \dots, \mathcal{X}_n^*$ ($n \geq 2$) zbioru $\mathbf{X}$ o następujących własnościach

$$\bigcup_{q=1}^n \mathcal{X}_q^* = \mathbf{X} \quad (2)$$

$$\mathcal{X}_r^* \cap \mathcal{X}_s^* = \{\text{zbiór pusty}\} \quad (s \neq r) \quad (3)$$

$$(x_i \in \mathcal{X}_r^*) \cap (x_j \in \mathcal{X}_s^*) \equiv \text{element } x_i \text{ jest preferowany w stosunku do } x_j, \quad (4)$$

w przypadku $r < s$,

$$x_i, x_j \in \mathcal{X}_r^* \equiv x_i, x_j - \text{elementy równoważne} \quad (5)$$

Relacja $\mathcal{X}_1^*, \dots, \mathcal{X}_n^*$ może być określona przy użyciu funkcji $T(x_i, x_j)$ $((x_i, x_j) \in \mathbf{X} \times \mathbf{X})$, o postaci

$$T(x_i, x_j) = r - s \quad (6)$$

gdzie: $x_i \in \mathcal{X}_r^*$ oraz $x_j \in \mathcal{X}_s^*$.

Wartość $T(x_i, x_j)$ nazywa się różnicą rang elementów x_i oraz x_j .

Postać relacji $\mathcal{X}_1^*, \dots, \mathcal{X}_n^*$ należy wyznaczyć (oszacować) na podstawie N -krotnych ($N \geq 1$) porównań parami elementów zbioru $\mathbf{X} \times \mathbf{X}$, określających różnice rang, zakłóconych błędami losowymi.

Dokładniej, zakłada się, że:

- k -te porównanie ($1 \leq k \leq N$) każdej pary $(x_i, x_j) \in \mathbf{X} \times \mathbf{X}$, oznaczane symbolem $g_k(x_i, x_j)$, stanowi ocenę wartości $T(x_i, x_j)$ zakłóconą błędem losowym losowym, tzn. jest zmienną losową, która może się różnić, z określonym prawdopodobieństwem, od wartości $T(x_i, x_j)$,

- porównanie $g_k(x_i, x_j)$ może przyjąć wartość ze zbioru $\{-(m-1), \dots, 0, \dots, m-1\}$, ponieważ nie zakłada się znajomości liczby podzbiorów n .

Przyjmuje się następujące założenia na temat rozkładów prawdopodobieństw błędów porównań:

Z1. Rozkład błędu porównania $g_k(x_i, x_j) - T(x_i, x_j)$ każdej pary $((x_i, x_j) \in \mathbf{X} \times \mathbf{X})$ jest jednomodalny z wartością modalną i medianą równą zeru; mediana spełnia warunki

$$\sum_{l \leq 0} P(g_k(x_i, x_j) - T(x_i, x_j) = l) > 1/2$$

$$\sum_{l \geq 0} P(g_k(x_i, x_j) - T(x_i, x_j) = l) > 1/2$$

Z2. Porównania $g_k(x_i, x_j)$ oraz $g_l(x_r, x_s)$ ($l \neq k$) są stochastycznie niezależne, tzn.

$$P((g_k(x_i, x_j) - T(x_i, x_j) = c_{kij}) \cap (g_l(x_r, x_s) - T(x_r, x_s) = c_{lrs})) = \\ P((g_k(x_i, x_j) - T(x_i, x_j) = c_{kij})P(g_l(x_r, x_s) - T(x_r, x_s) = c_{lrs})).$$

Z3. Wartość oczekiwana $E(g_k(x_i, x_j))$ ($1 \leq k \leq N; \langle i, j \rangle \in R_m$) dowolnego porównania nie zależy od pozostałych porównań $g_l(x_r, x_s)$ (para indeksów $\langle r, s \rangle$ różna od $\langle i, j \rangle$); porównania $g_1(x_i, x_j), \dots, g_N(x_i, x_j)$ ($\langle i, j \rangle \in R_m$) mają identyczne rozkłady.

Porównania $g_k(x_i, x_j)$ spełniające powyższe założenia mogą być otrzymane przy wykorzystaniu testów statystycznych, np. elementy zbioru $\mathbf{X}$ mogą być realizacjami niezależnych zmiennych losowych $Y_1, \dots, Y_n$, których wartości oczekiwane spełniają warunek $(E(Y_s) - E(Y_r))/(s - r) = 1$ ($s = 2, \dots, n; s > r$); porównania parami polegają wówczas na weryfikacji hipotezy dotyczącej różnicy wartości oczekiwanych każdej pary realizacji.

2. Postać estymatorów

Jak już wspomniano, sformułowano dwa rodzaje estymatorów relacji R , o własnościach (1)-(6). Pierwszy z nich minimalizuje sumę modułów różnic między postacią relacji, określoną przy wykorzystaniu funkcji $T(x_i, x_j)$, i porównaniami $g_k(x_i, x_j)$ ($k = 1, \dots, N; (x_i, x_j) \in \mathbf{X} \times \mathbf{X}$). Drugi estymator

minimalizuje sumę modułów różnic między postacią relacji, określoną w analogiczny sposób, i medianą $g^{(me)}(x_i, x_j)$ z porównań $\{g_1(x_i, x_j), \dots, g_N(x_i, x_j)\}$, tj. wartością środkową $g_{(\frac{N+1}{2})}(x_i, x_j)$ w uporządkowanym zbiorze porównań $\{g_{(1)}(x_i, x_j), \dots, g_{(N)}(x_i, x_j)\}$.

Ocena postaci relacji otrzymana przy użyciu pierwszego estymatora stanowi rozwiązanie zadania optymalizacji dyskretnej

$$\min_{\chi_1^{(i)}, \dots, \chi_{r_i}^{(i)} \in F_X} \left\{ \sum_{\langle i, j \rangle \in R_m} \sum_{k=1}^N |t^{(i)}(x_i, x_j) - g_k(x_i, x_j)| \right\} \quad (7)$$

gdzie:

F_X – zbiór rozwiązań dopuszczalnych (rodzina wszystkich relacji słabego porządku w zbiorze $\mathbf{X}$),

$\chi_1^{(i)}, \dots, \chi_{r_i}^{(i)}$ – i -ty element zbioru F_X ,

R_m – zbiór par indeksów o postaci: $R_m = \{\langle i, j \rangle \mid 1 \leq i, j \leq m; j > i\}$,

$t^{(i)}(x_i, x_j)$ – funkcja, określająca postać relacji $\chi_1^{(i)}, \dots, \chi_{r_i}^{(i)}$, zdefiniowana analogicznie do funkcji $T(x_i, x_j)$.

Ocena postaci relacji otrzymana przy użyciu drugiego estymatora stanowi rozwiązanie zadania optymalizacji dyskretnej

$$\min_{\chi_1^{(i)}, \dots, \chi_{r_i}^{(i)} \in F_X} \left\{ \sum_{\langle i, j \rangle \in R_m} |t^{(i)}(x_i, x_j) - g^{(me)}(x_i, x_j)| \right\} \quad (8)$$

gdzie:

$g^{(me)}(x_i, x_j)$ – wartość środkowa w zbiorze $\{g_{(1)}(x_i, x_j), \dots, g_{(N)}(x_i, x_j)\}$.

Oceny relacji otrzymane na podstawie zadań (7) oraz (8) będą oznaczane symbolami – odpowiednio: $\hat{\chi}_1, \dots, \hat{\chi}_{\hat{n}}$ oraz $\hat{\chi}_1, \dots, \hat{\chi}_{\hat{n}}$, a funkcje określające te oceny – symbolami: $\hat{T}(x_i, x_j)$ oraz $\hat{T}(x_i, x_j)$.

Oba zadania: (7) i (8) mogą mieć rozwiązanie wielokrotne. Zadanie (8) zawiera N -krotnie mniejszą liczbę zmiennych, niż zadanie (7).

Postać estymatorów $\hat{T}(x_i, x_j)$ oraz $\hat{T}(x_i, x_j)$ wynika z własności zmiennych losowych

$$W_N^* = \sum_{\langle i, j \rangle \in R_m} \sum_{k=1}^N |g_k(x_i, x_j) - T(x_i, x_j)| \quad (9)$$

$$W_N^{(me)*} = \sum_{\langle i, j \rangle \in R_m} |g^{(me)}(x_i, x_j) - T(x_i, x_j)| \quad (10)$$

odpowiadających relacji $\chi_1^*, \dots, \chi_n^*$.

W pracach [3; 4] wykazano między innymi, że:

- wartość oczekiwana zmiennej losowej (9) jest mniejsza, niż wartość oczekiwana dowolnej zmiennej $\tilde{W}_N = \sum_{\langle i, j \rangle \in R_m} \sum_{k=1}^N |\tilde{t}(x_i, x_j) - g_k(x_i, x_j)|$,

odpowiadającej relacji $\tilde{\chi}_1, \dots, \tilde{\chi}_{\tilde{n}}$ różnej od $\chi_1^*, \dots, \chi_n^*$,

- wartość oczekiwana zmiennej losowej (10), odpowiadającej relacji $\chi_1^*, \dots, \chi_n^*$, jest mniejsza, niż wartość oczekiwana zmiennej

$\tilde{W}_N^{(me)} = \sum_{\langle i, j \rangle \in R_m} |\tilde{t}(x_i, x_j) - g^{(me)}(x_i, x_j)|$, odpowiadającej dowolnej relacji

$\tilde{\chi}_1, \dots, \tilde{\chi}_{\tilde{n}}$ różnej od $\chi_1^*, \dots, \chi_n^*$,

- prawdopodobieństwa

$$P\left(\sum_{\langle i, j \rangle \in R_m} \sum_{k=1}^N |g_k(x_i, x_j) - T(x_i, x_j)| < \sum_{\langle i, j \rangle \in R_m} \sum_{k=1}^N |g_k(x_i, x_j) - \tilde{t}(x_i, x_j)|\right) \quad (11)$$

$$P\left(\sum_{\langle i, j \rangle \in R_m} |g_k(x_i, x_j) - T(x_i, x_j)| < |g_k(x_i, x_j) - \tilde{t}(x_i, x_j)|\right) \quad (12)$$

dążą do jedności, gdy $N \rightarrow \infty$.

Własności wartości oczekiwanych zmiennych losowych W_N^* , $W_N^{(me)*}$

i $\tilde{W}_N$, $\tilde{W}_N^{(me)}$ oraz zbieżność prawdopodobieństw (11)-(12) stanowią racjonalną podstawę do konstrukcji estymatorów (7), (8), ale nie określają dokładności uzyskanych ocen; mierniki ich dokładności określono przy wykorzystaniu metod symulacyjnych.

Błędy rozważanych estymatorów są wielowymiarowymi zmiennymi losowymi – odpowiednio $\{\hat{T}(x_i, x_j) - T(x_i, x_j); \langle i, j \rangle \in R_m\}$ oraz $\{\hat{T}(x_i, x_j) - T(x_i, x_j); \langle i, j \rangle \in R_m\}$. Taka postać błędów jest niedogodna do celów analitycznych, zwłaszcza w przypadku zbiorów $\mathbf{X}$ o dużej liczności. Dlatego też zbadano rozkłady częstości błędów skalarnych (zsumowanych) o postaci

$$\hat{\Delta} = \sum_{\langle i, j \rangle \in R_m} |\hat{T}(x_i, x_j) - T(x_i, x_j)| \quad (13)$$

$$\widehat{\Delta} = \sum_{\langle i, j \rangle \in R_m} |\widehat{T}(x_i, x_j) - T(x_i, x_j)| \quad (14)$$

Błąd skalarny zawiera informacje na temat rozmiaru różnicy między oceną i relacją $\chi_1^*, \dots, \chi_n^*$, nie określa jednoznacznie postaci błędu. Może być wykorzystany jako „zgrubny” miernik oceny dokładności estymacji.

3. Parametry badania symulacyjnego

Badanie symulacyjne skalarnych błędów estymacji (13) i (14) zostało wykonane na podstawie trzech postaci relacji $\chi_1^*, \dots, \chi_n^*$, o następujących parametrach.

1. Zbiór X zawierający 9 elementów,
2. Trzy postacie relacji:
 - z dziewięcioma podzbiorami: $\{x_1\}, \dots, \{x_9\}$ (9×9),
 - z sześcioma podzbiorami: $\{x_1\}, \{x_2, x_3\}, \{x_4\}, \{x_5, x_6\}, \{x_7\}, \{x_8, x_9\}$ (6×9),
 - z trzema podzbiorami: $\{x_1, x_2\}, \{x_3, x_4, x_5\}, \{x_6, x_7, x_8, x_9\}$ (3×9).
3. Funkcja prawdopodobieństwa błędów porównań o postaci

$$P(g_k(x_i, x_j) = T(x_i, x_j)) = \alpha_{ij}$$

$$P(g_k(x_i, x_j) - T(x_i, x_j) = -l) = (1 - \alpha_{ij}) / 2L_{ij}^{(d)}$$

$$(L_{ij}^{(d)} = T(x_i, x_j) + (m - 1); \quad l = 1, \dots, L_{ij}^{(d)})$$

$$P(g_k(x_i, x_j) - T(x_i, x_j) = l) = (1 - \alpha_{ij}) / 2L_{ij}^{(u)}$$

$$(L_{ij}^{(u)} = m - 1 - T(x_i, x_j); \quad l = 1, \dots, L_{ij}^{(u)})$$

zbadano trzy wartości α_{ij} : 0,3334; 0,4167; 0,5000 ($\langle i, j \rangle \in R_m$).

4. Liczba N porównań każdej pary: 1, 3, 5, 7, 9.

Łącznie zbadano 90 przypadków, tj. trzy relacje, trzy postacie rozkładu błędów porównań, pięć wartości liczby porównań N , przy wykorzystaniu dwóch estymatorów. Każdy przypadek symulowano 200-krotnie, przy użyciu generatora liczb losowych z rozkładu równomiernego. Zastosowana funkcja prawdopodobieństwa błędów porównań ma następujące cechy: suma prawdopodobieństw ujemnych błędów porównań jest równa sumie prawdopodobieństw

błędów dodatnich, błędy dodatnie mają równe prawdopodobieństwa, błędy ujemne mają równe prawdopodobieństwa. W przypadku $m = n$ występują również rozkłady jednostronne, tj. nieujemne (lub niedodatnie) – wówczas błędy dodatnie (lub ujemne) mają identyczne wartości.

4. Wyniki badania symulacyjnego

Wyniki badania symulacyjnego mają postać rozkładów częstości błędów (13) i (14). Częstości te określono w 10 przedziałach: pierwszy przedział – jednopunktowy – odnosi się do oceny bezbłędnej (błąd równy zeru), ostatni – do błędów o wartości powyżej 72, osiem „środkowych” – do przedziałów o postaci $(8v, 8(v+1)]$ ($v = 0, \dots, 8$). Rozpiętość „środkowych” przedziałów określono w taki sposób, aby zawierały oceny o podobnej odległości od rzeczywistej (symulowanej) relacji. Wyniki badania symulacyjnego zamieszczono w tabelach 1-9.

Jak już zaznaczono, zadania (7), (8) mogą generować oceny wielokrotne; w przypadku takim przyjęto błąd skalarny estymacji równy średniej arytmetycznej z błędów z poszczególnych ocen.

Tabela 1

Częstości błędów (13), (14); relacja: 9×9 oraz $\alpha = 0,3334$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \bar{\Delta}$	$\hat{\Delta}$	$\bar{\Delta}$	$\hat{\Delta}$	$\bar{\Delta}$	$\hat{\Delta}$	$\bar{\Delta}$	$\hat{\Delta}$	$\bar{\Delta}$
0	0,140	0,780	0,545	0,930	0,800	0,995	0,955	0,995	0,980
(0, 8]	0,035	0,075	0,175	0,035	0,100	0	0,020	0	0,010
(8, 16]	0,055	0,045	0,125	0,025	0,050	0	0,005	0,005	0,010
(16, 24]	0,040	0,040	0,050	0	0,030	0	0,015	0	0
(24, 32]	0,080	0,035	0,055	0	0,010	0,005	0,005	0	0
(32, 40]	0,120	0,005	0,045	0,005	0	0	0	0	0
(40, 48]	0,060	0,015	0,020	0	0,005	0	0	0	0
(48, 56]	0,105	0,005	0,020	0	0,005	0	0	0	0
(56, 64]	0,105	0	0,010	0,005	0	0	0	0	0
(64, 72]	0,080	0	0,005	0	0	0	0	0	0
>72	0,180	0	0	0	0	0	0	0	0

Tabela 2

Częstości błędów (13), (14); relacja 9×9 oraz $\alpha = 0,4167$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$
0	0,350	0,890	0,770	0,990	0,965	1,000	0,985	1,000	1,000
(0, 8]	0,090	0,065	0,075	0,005	0,015	0	0,010	0	0
(8, 16]	0,085	0,010	0,050	0	0,015	0	0,005	0	0
(16, 24]	0,100	0,015	0,040	0,005	0	0	0	0	0
(24, 32]	0,065	0,015	0,030	0	0,005	0	0	0	0
(32, 40]	0,035	0,005	0,025	0	0	0	0	0	0
(40, 48]	0,085	0	0,010	0	0	0	0	0	0
(48, 56]	0,050	0	0	0	0	0	0	0	0
(56, 64]	0,045	0	0	0	0	0	0	0	0
(64, 72]	0,035	0	0	0	0	0	0	0	0
>72	0,060	0	0	0	0	0	0	0	0

Tabela 3

Częstości błędów (13), (14); relacja 9×9 oraz $\alpha = 0,500$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$
0	0,600	0,985	0,920	1,000	0,995	1,000	1,000	1,000	1,000
(0, 8]	0,095	0,010	0,040	0	0,005	0	0	0	0
(8, 16]	0,080	0,005	0,015	0	0	0	0	0	0
(16, 24]	0,025	0	0,010	0	0	0	0	0	0
(24, 32]	0,050	0	0,010	0	0	0	0	0	0
(32, 40]	0,040	0	0,005	0	0	0	0	0	0
(40, 48]	0,025	0	0	0	0	0	0	0	0
(48, 56]	0,035	0	0	0	0	0	0	0	0
(56, 64]	0,015	0	0	0	0	0	0	0	0
(64, 72]	0,015	0	0	0	0	0	0	0	0
>72	0,025	0	0	0	0	0	0	0	0

Tabela 4

Częstości błędów (13), (14); relacja 6×9 oraz $\alpha = 0,3334$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$
0	0,025	0,420	0,155	0,775	0,485	0,915	0,635	0,990	0,825
(0, 8]	0,045	0,220	0,240	0,185	0,329	0,075	0,275	0,010	0,165
(8, 16]	0,100	0,165	0,245	0,015	0,125	0,005	0,070	0	0,005
(16, 24]	0,190	0,110	0,195	0,010	0,060	0,005	0,020	0	0,005
(24, 32]	0,195	0,040	0,055	0,015	0,005	0	0	0	0
(32, 40]	0,165	0,025	0,075	0	0,01	0	0	0	0
(40, 48]	0,090	0,015	0,020	0	0,005	0	0	0	0
(48, 56]	0,100	0	0,010	0	0	0	0	0	0
(56, 64]	0,025	0	0	0	0	0	0	0	0
(64, 72]	0,030	0,005	0	0	0	0	0	0	0
>72	0,035	0	0	0	0	0	0	0	0

Tabela 5

Częstości błędów (13), (14); relacja 6×9 oraz $\alpha = 0,4167$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$
0	0,115	0,790	0,560	0,970	0,815	0,995	0,945	1,000	0,970
(0, 8]	0,090	0,155	0,260	0,025	0,145	0,005	0,055	0	0,030
(8, 16]	0,170	0,020	0,120	0	0,030	0	0	0	0
(16, 24]	0,190	0,025	0,004	0,005	0,005	0	0	0	0
(24, 32]	0,150	0,010	0	0	0,005	0	0	0	0
(32, 40]	0,100	0	0,015	0	0,005	0	0	0	0
(40, 48]	0,090	0	0,005	0	0	0	0	0	0
(48, 56]	0,025	0	0	0	0	0	0	0	0
(56, 64]	0,035	0	0	0	0	0	0	0	0
(64, 72]	0,020	0	0	0	0	0	0	0	0
>72	0,015	0	0	0	0	0	0	0	0

Częstości błędów (13), (14); relacja 6×9 oraz $\alpha = 0,5000$

Tabela 7[illegible]

Tabela 8

Częstości błędów (13), (14); relacja 3×9 oraz $\alpha = 0,4167$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$
0	0,145	0,785	0,530	0,970	0,820	0,990	0,920	1,000	1,000
(0, 8]	0,255	0,190	0,360	0,030	0,165	0,010	0,080	0	0
(8, 16]	0,340	0,020	0,100	0	0,010	0	0	0	0
(16, 24]	0,170	0,005	0,010	0	0,005	0	0	0	0
(24, 32]	0,045	0	0	0	0	0	0	0	0
(32, 40]	0,020	0	0	0	0	0	0	0	0
(40, 48]	0,015	0	0	0	0	0	0	0	0
(48, 56]	0,005	0	0	0	0	0	0	0	0
(56, 64]	0,005	0	0	0	0	0	0	0	0
(64, 72]	0	0	0	0	0	0	0	0	0
>72	0	0	0	0	0	0	0	0	0

Tabela 9

Częstości błędów (13), (14); relacja 3×9 oraz $\alpha = 0,500$

Wartość błędu	$N = 1$	$N = 3$		$N = 5$		$N = 7$		$N = 9$	
	$\hat{\Delta} = \hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$	$\hat{\Delta}$
0	0,310	0,915	0,795	0,995	0,940	1,000	0,975	1,000	1,000
(0, 8]	0,325	0,085	0,190	0,005	0,060	0	0,025	0	0
(8, 16]	0,215	0	0,015	0	0	0	0	0	0
(16, 24]	0,130	0	0	0	0	0	0	0	0
(24, 32]	0,015	0	0	0	0	0	0	0	0
(32, 40]	0,005	0	0	0	0	0	0	0	0
(40, 48]	0	0	0	0	0	0	0	0	0
(48, 56]	0	0	0	0	0	0	0	0	0
(56, 64]	0	0	0	0	0	0	0	0	0
(64, 72]	0	0	0	0	0	0	0	0	0
>72	0	0	0	0	0	0	0	0	0

Zadania (7), (8), stanowiące podstawę do wyznaczenia ocen, rozwiązano metodą przeglądu pełnego, przy zastosowaniu liczby podzbiorów: $n - 1$, n , $n + 1$ (w przypadku $n = 9$, pominięto $n + 1$). Analiza wyników estymacji prowadzi do następujących wniosków.

1. Estymator oparty na rozwiązaniu zadania (7) zapewnia, w przypadku wielokrotnych porównań każdej pary ($N > 1$), lepszą precyzję estymacji niż estymator oparty na medianach z porównań każdej pary, tj. na rozwiązaniu zadania (8). Określenie większa precyzja oznacza: wyższą częstość oceny bezbłędnej oraz większą koncentrację rozkładu częstości w otoczeniu błędu równego zero. Różnica precyzji jest niewielka w przypadku rozkładów częstości obu estymatorów bliskich jednopunktowemu, tj. częstości błędów równych zero bliskich jedności.

2. Wzrost liczby porównań N powoduje szybkie zwiększanie precyzji estymacji. W przypadku: estymatora opartego na sumie różnic oraz prawdopodobieństwa bezbłędного porównania $\alpha = 0,5$, trzy porównania ($N = 3$) zapewniają ocenę relacji bliską bezbłędnej; w przypadku estymatora opartego na medianach liczba porównań N musi być większa o dwa. W przypadku pozostałych wartości α , tj. 0,3334 i 0,4167 oraz estymatora opartego na sumie różnic, liczba porównań, gwarantująca częstość bliską jedności, musi być nie mniejsza niż – odpowiednio: siedem i pięć. W przypadku estymatora opartego na medianach liczba ta musi być większa (co najmniej o dwa). Dokonanie 7-9 porównań każdej pary gwarantuje precyzję estymacji bliską lub równą 100%.

3. Najlepsza precyzja estymacji ma miejsce w przypadku relacji mocnego porządku, tj. $n = 9$; pozostałe przypadki, tj. $n = 3$ oraz $n = 6$, charakteryzują się nieco niższą – podobną precyzją. W przypadku relacji mocnego porządku akceptowalną precyzję (częstość bezbłędного porównania przekraczającą 50%) zapewnia pojedyncze porównanie każdej pary $N = 1$ oraz $\alpha = 0,5$.

4. Rozkłady częstości błędów estymacji charakteryzują się dużą koncentracją w przypadku, gdy częstość bezbłędnej oceny postaci relacji przekracza 0,5; częstość błędu nie przekraczającego wartości 16 jest wówczas bliska jedności. Jeśli częstość bezbłędnej oceny przekracza 0,75, to analogiczna własność dotyczy błędu estymacji nie przekraczającego 8. Mały błąd estymacji implikuje nieznaczną różnicę między oceną i rzeczywistą relacją.

5. W przypadku, gdy nie wszystkie rozkłady porównań parami spełniają założenia Z1, że zero jest medianą i dominantą w rozkładzie porównań parami, wyniki estymacji stają się nieakceptowalne; sytuacja taka ma miejsce w przypadku: $N = 1$ i $\alpha \leq 0,4167$ oraz $N \leq 3$ i $\alpha = 0,3334$.

Powyższa analiza wyników symulacji stanowi podstawę do określenia: liczby porównań N oraz wartości prawdopodobieństwa α niezbędnych do uzyskania wymaganej precyzji estymacji. Należy podkreślić, że rozkład błędu porównania każdej pary, zastosowany w badaniu symulacyjnym (o identycznych prawdopodobieństwach w lewym i w prawym ogonie rozkładu), jest niekorzystny z punktu widzenia precyzji estymacji; wykorzystanie rozkładu ściśle monotonicznego, w którym większym (co do modułu) błędom odpowiadają mniejsze prawdopodobieństwa, zapewnia polepszenie precyzji ocen.

Podsumowanie

Wyniki badań symulacyjnych, omówionych w niniejszej pracy, są podstawą do następujących – syntetycznych wniosków:

- estymatory relacji porządku oparte na porównaniach w postaci różnicy rang, charakteryzują się dobrą precyzją, przy słabych założeniach na temat błędów porównań,
- najlepszą precyzję estymacji uzyskuje się w przypadku relacji mocnego porządku,
- estymator oparty na sumie różnic między funkcją określającą postać relacji a porównaniami wykazuje istotną przewagę precyzji nad estymatorem opartym na medianach z porównań,
- precyzja estymacji polepsza się znacząco wraz z liczbą porównań N ; można wyznaczyć liczebność porównań zapewniającą ocenę bezbłędną lub o nieznacznym błędzie.

Można postulować, na podstawie symulacji dokonanych na zbiorach 10-12-elementowych, że powyższe wnioski zachowują słuszność w przypadku estymacji relacji R w zbiorach o większej liczebności. Należy dodać, że wyznaczenie ocen relacji w przypadku $n > 12$ wymaga stosowania algorytmów optymalizacji dyskretniej lub algorytmów przybliżonych.

Literatura

1. David H.A. (1988). The Method of Paired Comparisons. Ch. Griffin, London.
2. Klukowski L. (1994). Some Probabilistic Properties of the Nearest Adjoining order Method and its Extensions. *Annals of Operational Research*, 51, s. 241-261.
3. Klukowski L. (2007). Estimation of the Preference Relation on the Basis of Medians from Pairwise Comparisons in the Form of Difference of Ranks. W: *Computer Recognition Systems 2*. Red. Kurzynski (et al.). Springer, Berlin, Heidelberg, New York.
4. Klukowski L. (2008). Estimation of the Preference Relation on the Basis of Multiple Pairwise Comparisons in the Form of Difference of Ranks. *Control and Cybernetics*, 37, 3, s. 711-729.
5. Klukowski L. (2008). Estymacja relacji porządku na podstawie porównań parami z błędami losowymi. W: *Modelowanie preferencji a ryzyko '09*. Red. T. Trzaskalik. AE, Katowice, s. 61-76.
6. Slater P. (1961). Inconsistencies in a Schedule of Paired Comparisons. *Biometrika*, 48, s. 303-312.

**PROPERTIES OF ESTIMATORS OF THE PREFERENCE RELATION, BASED
ON DIFFERENCES OF RANKS – SIMULATION STUDY**

Summary

In the papers of the author (see e.g. Klukowski (2008a, 2008b) have been proposed the estimators of the preference relation in a finite set on the basis of multiple pairwise comparisons, in the form of differences of ranks, disturbed by random errors. Two estimators, based on the idea nearest adjoining order have been considered: they minimize differences between relation form and comparisons. The first estimator is based on total number of differences, the second – on the sum comprising medians from comparisons of each pair. The properties of both estimators, obtained in analytical way, do not comprise distributions of comparisons' errors. In this work the results of simulation survey on this subject are presented. They indicate high efficiency of the estimators proposed. Moreover, they allow determining the number of comparisons providing necessary precision of estimates for given distributions of comparisons' errors.

Donata Kopańska-Bródka

Uniwersytet Ekonomiczny w Katowicach

WZORZEC MODELU OCZEKIWANEJ UŻYTECZNOŚCI A ROZWÓJ TEORII DECYZJI

Wprowadzenie

Koncepcja modelu teorii oczekiwanej użyteczności po raz pierwszy została zaproponowana w 1738 roku przez Daniela Bernoulliego jako odpowiedź na pytanie, ile rozsądna osoba powinna zapłacić za udział w grze, której oczekiwana wygrana w jednostkach monetarnych jest nieskończenie wielka. Problem ten nazywany „paradoksem petersburskim” polegał na tym, że w sytuacji, kiedy obowiązującą oceną gry losowej była jej wartość oczekiwana to, pomimo nieskończenia dużej oczekiwanej wygranej, większość ludzi za udział w takiej grze zapłaciłaby niewielką kwotę. Bernoulli zauważył, że dla jednostki oczekiwana wartość monetarna gra nie jest postrzegana jako „wartość” gry. Twierdził, że ludzie w sposób subiektywny przypisują wartościom monetarnym inne wielkości nazywane użytecznością, zaś wartością gry jest oczekiwana użyteczność (wtedy to nazwano wartością moralną). Zaproponował logarytmiczną transformację wartości monetarnych, za pomocą której rozwiązał problem gry petersburskiej. Stąd też pierwszą analityczną postacią funkcji użyteczności była funkcja logarytmiczna. Teorię Bernoulliego nie zajmowano się aż do 1947 roku, w którym John von Neuman i Oscar Morgenstern pokazali, że hipotezę oczekiwanej użyteczności można wykazać przyjmując określony zbiór aksjomatów dotyczących własności relacji preferencji. Aksjomaty te uznano za zasady racjonalnego wyboru, które akceptuje każda rozsądna jednostka. Zasada oczekiwanej użyteczności stała się normatywną zasadą wyboru w sytuacji ryzyka, a także wskazaniem (praktyczną poradą), jak ludzie powinni wybierać. Chociaż teoria oczekiwanej użyteczności postrzegana jest jako narzędzie modelowania decyzji zgodne z powszechną intuicją, to zachowania decydentów obserwowane w rzeczywistych sytuacjach podają w wątpliwość normatywny charakter tej teorii. Zaobserwowane paradoksy i niezgodności stały się dla naukowców impulsem do dalszych badań, a tym samym do rozszerzenia teorii. Szczególną uwagę zwrócono na wagi decyzyjne przypisy-

wane poszczególnym konsekwencjom (wynikom) decyzji. Przeprowadzane eksperymenty i obserwacje potwierdziły tezę o dokonywaniu przez jednostki subiektywnej transformacji prawdopodobieństw. Rozszerzenia teorii oczekiwanej użyteczności zaproponowane między innymi przez Karmarkara (1978), Quiggina (1982), Yaary'ego (1987) i Machiny (1989) przyjęto nazywać teoriami konwencjonalnymi. Teorie te badają własności relacji preferencji potrzebne do wyjaśnienia źródła niezgodności z aksjomatem niezależności. W tych rozszerzeniach przyjmuje się aksjomatykę teorii oczekiwanej użyteczności, jednak funkcjonal określany na zbiorze ryzykownych decyzji nie jest wartością oczekiwaną, lecz transformacją użyteczności możliwych wyników ważoną wagami decyzyjnymi. Ta nowa zasada nie prowadziła już do takich niezgodności jak zasada maksymalizacji wartości oczekiwanej.

Teorie konwencjonalne, mimo znacznego poprawienia zgodności z aksjomatyką wprowadzoną przez von Neumana i Morgensterna, nie radziły sobie z wytłumaczeniem obserwowanego w eksperymentach „paradoksu odwrócenia”, polegającego na zależności wyboru decyzji od sposobu opisu rozkładu*. Odpowiedzią na problemy i niezgodności teorii konwencjonalnych zaobserwowane w eksperymentach i praktyce decyzyjnej był nowy nurt, określany teoriami alternatywnymi bądź niekonwencjonalnymi. U podstaw teorii niekonwencjonalnych (non-expected theory) leży założenie, że relacji preferencji jednostki nie można modelować tylko i wyłącznie za pomocą czystych funkcyjnych zależności w sztywnych ramach jakie stwarzają zakładane a priori aksjomaty. Kahneman i Tversky (1979) sformułowali aksjomatykę nowej teorii zgodną z obserwowanymi zachowaniami jednostek w obliczu ryzykownych wyborów. Nurt ten, nazywany teorią perspektywy, modeluje zachowania jednostek w sytuacji wyboru. Zaobserwowane niezgodności tej teorii z zasadą dominacji stochastycznej pozwoliły Tversky'emu i Kahnemanowi (1992) dokonać jej rozszerzenia nazywanego kumulacyjną teorią perspektywy. Bardzo szerokiej analizy teorii konwencjonalnych i niekonwencjonalnych dokonał Starmer [9] w 2000 roku w pracy, do której odnosi się większość badaczy problemów wyboru.

1. Teoria Oczekiwanej Użyteczności – wzorcowa metodologia

Dla dalszych rozważań przyjmijmy oznaczać zbiór możliwych wariantów decyzyjnych symbolem $\mathbb{X}$, zbiór możliwych konsekwencji (zdarzeń, wyników) symbolem $\mathbb{W}$, dwuargumentową relację preferencji symbolem $\succeq$ oraz relację nierozróżnialności (indyferencji) symbolem $\approx$. Funkcja $u(\cdot)$ jest funkcyjną

* W literaturze paradoks ten jest znany pod nazwą „choroby azjatyckiej”.

reprezentacją relacji preferencji w zbiorze wyników, natomiast funkcyjną reprezentacją preferencji w zbiorze wariantów $\mathbb{X}$ oznaczana jest funkcjonałem $V(\cdot)$. Zbiór możliwych wyników dla wariantu X jest wektorem $(w_1, w_2, \dots, w_n)$ gdzie $w_i \in \mathbb{W}$ oraz p_i jest prawdopodobieństwem uzyskania wyniku w_i . Dla każdego $X \in \mathbb{X}$ przyjmujemy oznaczenie

$$X = \left\{ (w_1, p_1), (w_2, p_2), \dots, (w_n, p_n) \quad \text{oraz} \quad \sum_{i=1}^n p_i = 1 \right\}$$

W teorii oczekiwanej użyteczności funkcyjną reprezentację $u(\cdot)$ dwuargumentowej relacji preferencji w zbiorze wyników ($u: \mathbb{W} \rightarrow \mathbb{R}$) nazywa się funkcją użyteczności i jest to funkcja niemalejąca. Ponadto przyjmuje się, że dwuargumentowa relacja preferencji w zbiorze wariantów decyzyjnych spełnia następujące aksjomaty zupełności, przechodności, ciągłości oraz niezależności

A1. Zupełność (spójność)

$$(X \succeq Y) \vee (Y \succeq X) \vee (X \approx Y)^*$$

A2. Przechodność

$$(X \succeq Y) \wedge (Y \succeq Z) \Rightarrow (X \succeq Z)$$

A3. Ciągłość

Jeśli $X \succeq Y \succeq Z$ to istnieje takie $p \in (0, 1)$, że skonstruowane loterie $L_1 = \{(X, p); (Z, 1-p)\}$ i $L_2 = \{(Y, 1)\}$ są jednakowo oceniane (indyferentne), czyli $L_1 \approx L_2$

A4. Niezależność

Dla dowolnych wariantów X, Y i Z jeśli $X \succeq Y$, to dla każdego $p \in (0, 1)$ loterie $L_1 = \{(X, p); (Z, 1-p)\}$ i $L_2 = \{(Y, p); (Z, 1-p)\}$ spełniają $L_1 \succeq L_2$.

Relacja spełniająca aksjomat A1 nazywany postulatem rozstrzygalności oraz A2 jest relacją słabego porządku w zbiorze wariantów decyzyjnych. Jeśli natomiast spełnione są A1-A3 to istnieje reprezentacja funkcyjna relacji $V: \mathbb{X} \rightarrow \mathbb{R}$ taka, że

$$X \succeq Y \Leftrightarrow V(X) \geq V(Y)$$

i mówimy wówczas, że funkcja V jest zgodna z relacją $\succeq$.

Aksjomat niezależności A4 jest bardzo silnym aksjomatem, jednak istotnie uściśla relację preferencji oraz pozwala nadać teorii wymiar praktyczny. Jeśli spełnione są aksjomaty A1-A4, to funkcja

$$V(X) = \sum_i u(w_i) \cdot p_i \quad (1)$$

* Relacja indyferencji $\approx$ definiowana jest $X \approx Y \Leftrightarrow (X \succeq Y) \wedge (Y \succeq X)$.

jest funkcyjną reprezentacją relacji w zbiorze wariantów. Formuła (1) jest oczekiwaną użytecznością wariantu decyzyjnego zgodną z preferencjami jednostki spełniającymi przyjęte aksjomaty.

Teoria oczekiwanej użyteczności w prosty sposób określa zasadę wyboru ryzykownych wariantów decyzyjnych spełniającą warunek zgodności z zasadą dominacji stochastycznej rzędu pierwszego (FSD). W jasny sposób interpretuje awersję jednostki do ryzyka, jako preferowanie wyboru pewnej wypłaty w niż jakiegokolwiek loterii X o wartości oczekiwanej $E(X) = w$. Natomiast skłonność do ryzyka to preferowanie wyboru dowolnej loterii X o wartości oczekiwanej $E(X) = w$ niż pewnego otrzymania wypłaty w . Ponadto, krzywizna wykresu funkcji użyteczności jest jedynym i całkowitym nośnikiem informacji o stosunku decydenta do ryzyka. Niezaprzeczalnym walorem teorii EU (oczekiwanej użyteczności) jest jej jasna i przejrzysta aksjomatyka, duża ogólność oraz możliwość stosowania miar Pratt-Arrowa* jako narzędzi do mierzenia siły awersji do ryzyka. Ponadto, formalna analiza obserwowanych ważnych ekonomicznych zachowań jest możliwa i akceptowalna w ramach tak rygorystycznej teorii. Problematyka wyboru optymalnego portfela akcji jest szczególnie przykładem praktycznej aplikacji teorii EU.

Od momentu sformułowania tej aksjomatycznej teorii trwa nieustanna dyskusja nad jej zgodnością z praktyką podejmowania decyzji oraz obserwowanym zachowaniem jednostek w sytuacji konieczności dokonywania wyboru. Problemy, które są w konflikcie z teorią EU można ująć w następujących trzech grupach zagadnień [8]:

- 1) zaobserwowana w eksperymentach (paradoks Allais, paradoks Ellsberga) niezgodność z aksjomatem niezależności preferencji,
- 2) konstrukcja funkcji użyteczności oparta na prostym kwestionariuszu,
- 3) obserwowane zachowania jednostek, polegające na potrzebie ubezpieczenia od ryzyka przy jednoczesnym udziale w grach i loteriach (olbrzymie zapotrzebowanie na bilety i kupony loteryjne).

Próby wyeliminowania powyższych problemów przyczyniły się do powstania wielu nowych koncepcji w teorii podejmowania decyzji. Formalne sposoby eliminacji występującej niezgodności z aksjomatem niezależności leżą u podstaw wszystkich teorii stanowiących rozszerzenie bądź uogólnienie teorii oczekiwanej użyteczności.

Powstałe teorie przyjmują matematyczną strukturę formuły zasady oczekiwanej użyteczności, ale w większości zastępują ją liniowym funkcjonałem, w którym prawdopodobieństwa otrzymania poszczególnych wyników zastą-

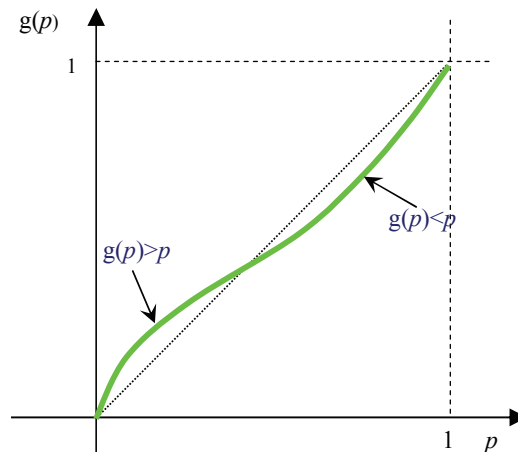
* Miary bezwzględnej i względnej awersji do ryzyka.

pione są wagami decyzyjnymi związanymi z danym wynikiem. Konstruowany jest nowy funkcjonal, który zapewnia zgodność z aksjomatem niezależności. Zatem wzorcem funkcjonalu jest liniowa kombinacja ocen wyników (użyteczności) i wag decyzyjnych. Z innych znanych i uznawanych konstrukcji, odbiegających od tego wzorca jest propozycja Machiny (1989), w której występuje zależność kwadratowa. W literaturze można znaleźć różne formuły subiektywnych transformacji prawdopodobieństw i określania wag decyzyjnych. Nieliniowe transformacje prawdopodobieństwa nazywane są funkcjami wag i w dalszej części oznaczane będą symbolem $g(p)$, natomiast symbolem π_i oznaczana będzie waga decyzyjna związana z wynikiem o numerze i . Wartości funkcji $g(p)$ są subiektywnymi ocenami obiektywnych prawdopodobieństw p .

Funkcja wag $g(\cdot)$ jest ciągłą, nieliniową i rosnącą transformacją typu $g: [0, 1] \rightarrow [0, 1]$, której wykres ma kształt odwróconego S, $g(0) = 0$ oraz $g(1) = 1$. Bezpośrednie zastępowanie prawdopodobieństw ich subiektywnymi transformacjami w formule oczekiwanej użyteczności (1) może wyjaśniać paradoks Allais, ale takie rozszerzenie teorii EU nie jest akceptowalne, ponieważ zostaje naruszona zasada zgodności z dominacją stochastyczną rzędu pierwszego*. Zatem przyjmowanie wag decyzyjnych na poziomie $g(p_i)$ powoduje, że odpowiedni funkcjonal jako reprezentacja preferencji w zbiorze decyzji prowadzi do teorii, które nie mogą być akceptowane ani jako normatywny model, ani jako deskryptywny model wyboru.

W literaturze proponuje się modyfikacje wag decyzyjnych, które zależą zarówno od subiektywnych transformacji poszczególnych prawdopodobieństw, jak i pełnego rozkładu prawdopodobieństwa wariantu decyzyjnego. W teoriach konwencjonalnych szczególnie popularne są nieliniowe transformacje kumulowanych prawdopodobieństw czy też złożenie funkcji wag z kumulacyjną dystrybuantą lub dekumulacyjną dystrybuantą rozkładu prawdopodobieństwa. Formalny problem zgodności oceny relacji preferencji z dominacjami stochastycznymi oraz dowody stosownych twierdzenia bardzo szeroko omawia Fishburn (1978). Wykazał on, że dla loterii prostych (rozkładów dwupunktowych) zastąpienie w formule oczekiwanej użyteczności prawdopodobieństw p_i ich subiektywnymi ocenami $g(p_i)$ nie narusza zasady FSD tylko wówczas, jeśli $g(p_i) = p_i$. Na rys. 1 przedstawiono wykres funkcji wag $g(p)$ o krzywiźnie typowej dla rozszerzonych teorii oczekiwanej użyteczności.

* Czasami mowa o warunku (aksjomacie) monotoniczności (monotonicity) preferencji ze względu na rosnące prawdopodobieństwo.



Rys. 1. Funkcja wag

Poprzez funkcję wag małe prawdopodobieństwa są zawyżane natomiast duże niedoszacowanie. Jeśli funkcja wag określona jest dla prawdopodobieństw skumulowanych, to przeszacowane są wyniki w lewym ogonie rozkładu, jak i w prawym ogonie rozkładu. Zatem zawyżanie prawdopodobieństw zdarzeń „gorszych” niż dane zdarzenie wyraża dużą awersję do ryzyka i pesymizm jednostki (skłonność do ubezpieczenia się od zdarzeń o niskim prawdopodobieństwie), zaś zawyżanie „lepszyc” niż dane zdarzenie charakteryzuje optymizm jednostki, a co za tym idzie udział w loteriach i grach losowych (popyt na loterie). Modelowanie postaci funkcji wag to bardzo szeroka możliwość funkcyjnych reprezentacji indywidualnych preferencji.

2. Konwencjonalne teorie oczekiwanej użyteczności

W konwencjonalnych rozszerzeniach teorii oczekiwanej użyteczności przyjmuje się założenie, że w percepcji decydenta obiektywne prawdopodobieństwa podlegają zniekształceniom zanim będą użyte jako dane wejściowe w procesie decyzyjnym. Użyteczność wyniku zależy nie tylko od prawdopodobieństwa jego uzyskania, ale także od jego pozycji rankingowej względem pozostałych możliwych wyników. Stąd też przyjęto nazywać to podejście RDEU, co jest skrótem pochodzącym od nazwy angielskiej teorii oczekiwanej użyteczności zależnej od rankingu (Rank Dependent Expected Utility). Zatem podejście to uwzględnia jak „dobre” bądź „złe” są poszczególne wyniki w porównaniu z innymi.

Modele RDEU to takie modele decyzyjne, w których wagi decyzyjne $\pi_i(p)$ są nieliniową, niemalejącą transformacją $g(\cdot)$ określoną na wartościach funkcji dystrybucyjnej. Twórcą takiego podejścia jest Quiggin [8], który pokazał, że jego model mieści się w aksjomatyce teorii oczekiwanej użyteczności, ale ma tę przewagę, że nie narusza aksjomatu niezależności, natomiast model EU jest szczególnym przypadkiem modelu RDEU. Zaproponowany funkcjonal wartościujący losowe warianty decyzyjne $X \in \mathbb{X}$ jest następującej postaci

$$V(X) = \sum_i u(w_i) \cdot \pi_i(p) \quad (2)$$

gdzie

$$\pi_i(p) = g\left(\sum_{k=1}^i p_k\right) - g\left(\sum_{k=1}^{i-1} p_k\right) \quad (3)$$

Występująca we wzorze (2) funkcja $u(w_i)$ jest funkcją użyteczności w sensie von Neumana i Morgensterna, natomiast funkcja wag $g(\cdot)$ jest ciągłą, rosnącą funkcją której wykres ma kształt odwróconego S. Wagi decyzyjne określone zależnością (3) dla takich samych wyników występujących w różnych wariantach decyzyjnych mogą się nawet istotnie różnić. Teoria ta nie stosuje się do wyborów o wynikach jednakowo prawdopodobnych.

Przyjmując, że funkcja dystrybucyjnej rozkładu zmiennej X jest oznaczana symbolem $F_X(t)$, zależność (3) można również przedstawić w postaci

$$\pi_i(p) = g(F(x_i)) - g(F(x_{i-1})) \quad (4)$$

Powyższe określenie wag decyzyjnych przeważnie jest używane w sytuacjach, kiedy wykorzystywany jest operator całkowania oraz wówczas, gdy ryzykowne warianty decyzyjne są zmiennymi losowymi o ciągłych rozkładach prawdopodobieństwa. Należy podkreślić, że w teorii RDEU nawet bardzo małe zmiany wartości poszczególnych wyników mogą mieć duży wpływ na ocenę decyzji, jeśli wpływają na zmianę pozycji rankingowej (wagi decyzyjne wzrastają) natomiast zmiany wartości wyników nawet bardzo duże nie wpływają na wagi decyzyjne jeśli nie naruszają uporządkowania

Zasada wyboru określona zależnością (2) przez wielu teoretyków i praktyków jest akceptowana, jednak nie rozwiązuje wszystkich paradoksów towarzyszącym teorii EU, a w szczególności paradoksu „choroby azjatyckiej”. Paradoks ten zaprezentowany przez Tversky’ego i Kahnemana (1981) Kahneman i Tversky (1981) polegał na decyzji dotyczącej wyboru programu walki z chorobą, która może spowodować śmierć 600 osób. Eksperti przedstawiają następujące dwie wersje programów walki z chorobą:

I wersja zawiera dwa programy A i B:

A – uratować można 200 osób z pewnością,

B – wszyscy zginą z prawdopodobieństwem 2/3.

II wersja zawiera programy C i D:

C – poniesie śmierć 400 osób z pewnością,

D – wszyscy przeżyją z prawdopodobieństwem $1/3$.

Jeśli prezentowana była wersja I, to większość preferowała program A, natomiast w wersji II preferowano program D. Paradoks polega na tym, że programy wersji I są stochastycznie identyczne jak programy wersji II, jednak przy dla różnych sformułowań problemu wyboru jednostki dokonywały przeciwnych wyborów tzn. $A \succeq B$ i $D \succeq C$.

Innym, również ważnym rozszerzeniem teorii oczekiwanej użyteczności jest zaproponowana przez Yaari'ego (1987) teoria dualna. W swoich koncepcjach Quiggin i Yaari są zgodni co do założenia, że w umyśle decydenta możliwe są zniekształcenia prawdopodobieństw wyników, zanim staną się one danymi wejściowymi w procesie decyzyjnym oraz że istnieje funkcja rzeczywista wartościująca te zniekształcenia i konieczna jest taka modyfikacja modelu EU, która uwzględni takie subiektywne postrzeganie prawdopodobieństw.

W standardowej teorii EU stosunek do ryzyka, jak i subiektywna ocena wyniku przez decydenta traktowane są łącznie w funkcji użyteczności. Mówi się wręcz, że fundamentalna zasada EU o malejącej marginalnej użyteczności wyniku, a awersja do ryzyka to synonimy. Yaari w swoim podejściu pokazuje, że te dwie cechy traktowane oddzielnie zachowują zgodność z kardynalną użytecznością. Podstawą konstrukcji jego funkcjonału wartościującego losowe warianty wyboru nie jest funkcja dystrybucyjna, jak to miało miejsce w podejściu Quiggina, ale dekulacyjny funkcja rozkładu prawdopodobieństwa oznaczana symbolem $G_X(t) = \Pr\{X > t\}$. W zasadach teorii dualnej przyjmuje się aksjomat neutralności*, aksjomaty relacji słabego porządku, niezależności oraz aksjomat monotoniczności gwarantujący zgodność z dominacją stochastyczną z tą różnicą, że odnoszą się do dekulacyjnych funkcji rozkładu $G(t)$, a tak sformułowany aksjomat niezależności nazywany jest niezależnością dualną. Yaari wykazał, że jeśli relacja preferencji spełnia zaproponowaną aksjomatykę, to następujący funkcjonal

$$V(X) = \int_0^1 g(G_X(t))dt \quad (5)$$

jest zgodny z relacją preferencji, czyli

$$X \succeq Y \Leftrightarrow \int_0^1 g(G_X(t))dt \geq \int_0^1 g(G_Y(t))dt$$

* Dwa losowe warianty są indyferentne, jeśli mają takie same dekulacyjne funkcje rozkładu prawdopodobieństwa.

Teoria dualna Yaari'ego z jednej strony wyjaśnia obserwowane niezgodności i paradoksy teorii oczekiwanej użyteczności, ale jednocześnie z drugiej generuje takie niezgodności, które z kolei wyjaśnia teoria EU. W tym sensie może być rozumiana dualność teorii. Dla przykładu, w teorii EU nie istnieje funkcja użyteczności, która tak samo porządkuje warianty jak współczynnik Giniego, natomiast w teorii dualnej taka zależność istnieje.

W teorii dualnej wszystkie losowe warianty przyjmują wartości z przedziału $[0, 1]$. Założenie to w żadnym wypadku nie umniejsza ogólności rozważań, ponieważ wystarczy realizacje losowych wariantów znormalizować. Warto zauważyć, że chociaż wartości funkcjonu (5) nie zależą od funkcji użyteczności von Neumana i Morgensterna określonej w zbiorze możliwych wyników, to jednak związek z zasadą maksymalnej oczekiwanej użyteczności istnieje i nie prowadzi do sprzeczności. Jeśli $u(x)$ jest funkcją użyteczności, to oczekiwana użyteczność jest równa

$$V(X) = E(u(X)) = -\int_0^1 u(t) dG_X(t) = \int_0^1 u(G_X^{-1}(p)) dp$$

co już wyraźniej pokazuje związek z (5).

Warto również wspomnieć o koncepcji uogólnienia EU zaproponowanej przez Marka Machiny w 1982 roku. Obserwacja Friedmana i Savage'a (1948), że niezależnie od poziomu dochodu jednostki wykazują wzrastającą wolę do kupowania ubezpieczenia, jak i wzrastające zapotrzebowanie na loterie i udział w grach hazardowych zainspirowała Machinę do prac nad wyjaśnieniem tego fenomenu w ramach teorii oczekiwanej użyteczności. Adekwatność proponowanej przez nich funkcji użyteczności z punktami przegięcia* została potwierdzona przez Machinę w jego koncepcji uogólnionej teorii oczekiwanej użyteczności. Zaproponowana przez niego aksjomatyka nie zawiera aksjomatu niezależności teorii EU, jak również aksjomatu dualnej niezależności Yaari'ego. W zamian proponuje aksjomat o istnieniu funkcjonu różniczkowalnego w sensie Frecheta. Machina zaproponował odpowiedni funkcjonal, stanowiący reprezentację funkcyjną relacji preferencji spełniającej założoną przez niego aksjomatykę. Formuła matematyczna tej zależności jest inna niż w pozostałych teoriach, ponieważ nie jest już sumą iloczynów użyteczności wyników i odpowiadających im wag decyzyjnych.

* Markowitz (1952) również uzasadniał, że jednostki zmieniają preferencje do ryzyka i funkcje użyteczności są typu S-shape. W swojej pracy sugeruje, że zmiany w stanie posiadania jednostki są przyczyną zmiany krzywizny funkcji użyteczności, zaś punkt przegięcia to bieżący lub zbliżony poziom stanu posiadania.

3. Teorie niekonwencjonalne

W omawianych wcześniej konwencjonalnych teoriach oczekiwanej użyteczności problem wyboru sprowadzany był do zagadnienia maksymalizacji preferencji. Metodologia taka, chociaż nie uwzględniała obserwowanych zachowań decyzyjnych i procesu prowadzącego do wyboru, zdominowała obowiązujące teorie ekonomiczne. Teorie, które uwzględniają psychologiczne aspekty prowadzące do wyboru nazwane zostały przez Starmera [9] teoriami proceduralnymi. Główną cechą takich teorii jest założenie, że jednostki dokonujące wyboru opierają się na swoistych heurystykach decyzyjnych. Zatem podstawowym zagadnieniem jest identyfikacja zbioru heurystyk, a następnie określenie warunków stosowania poszczególnych zasad. Dla psychologów modele oparte na takich złożeniach są uznawane i właściwą metodologią.

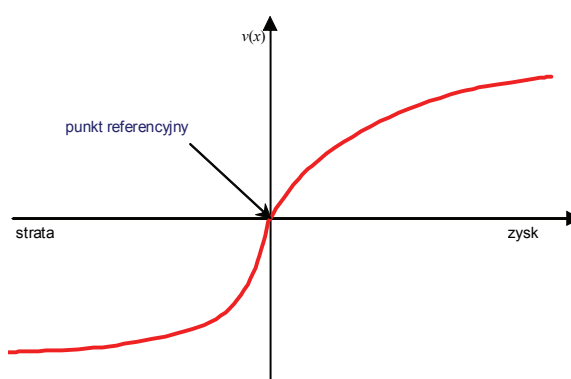
Zaproponowana przez psychologów, Daniela Kahnemana i Amosa Tversky'ego [5], teoria perspektywy (PT) od dwóch dekad należy do metodologii bardzo szeroko dyskutowanych zarówno wśród psychologów, jak i ekonomistów. W teorii perspektywy wybór wariantu ryzykownego jest procesem dwufazowym. W pierwszej fazie „edycji” warianty decyzyjne są korygowane przy użyciu różnych heurystyk, w fazie drugiej wariantom są przyporządkowane odpowiednie wartości funkcji preferencji, której forma matematyczna jest typu średniej użyteczności z wagami decyzyjnymi. Zwolennicy teorii wierzą, że proces taki jest zgodny z rzeczywistym zachowaniem jednostki w sytuacji wyboru. Istota tego podejścia polega na wyodrębnieniu fazy edycji oraz na relatywizmie wyników. Znaczący to, że konsekwencje decyzji interpretowane są jako straty bądź zyski w odniesieniu do pewnego punktu referencyjnego czy też status quo. Zatem decydent nie poświęca istotnej uwagi możliwym do uzyskania wynikom, lecz odchyleniom od jego indywidualnego punktu odniesienia. Dla skończonego zbioru wyników $W = \{w_1, w_2, \dots, w_n\}$ losowy wariant decyzyjny w dziedzinie wartości absolutnych definiowany jest w następujący sposób

$$X = \left\{ (w_i, p_i); \quad w_i \in W \text{ dla } i = 1, 2, \dots, n \quad \text{oraz} \quad \sum_{i=1}^n p_i = 1 \right\}$$

gdzie p_i jest prawdopodobieństwem wystąpienia realizacji w_i . Dla punktu referencyjnego w wariant daje perspektywę zysku $x_i = w_i - w$ jeśli $w_i > w$ lub perspektywę strat jeśli $w_i < w$. Zatem wariant X w dziedzinie zysków i strat (oznaczanych x) względem punktu odniesienia w bywa nazywany perspektywą (prospect) i oznaczany $X_w = \{(x_1, p_1), (x_2, p_2), \dots, (x_n, p_n)\}$.

Ponadto, zyski i straty oceniane są w odmienny sposób za pomocą funkcji wartości. Przez niektórych autorów funkcja wartości nazywana jest użytecznością, jednak sami twórcy teorii perspektywy konsekwentnie unikają takiej nazwy. Jedną z przyczyn może być idea niezależnego od ryzyka wartościowania zysków i strat, w przeciwieństwie do funkcji użyteczności teorii EU, która to wyraża również stosunek jednostki do ryzyka.

Teoria perspektywy opiera się na dwóch podstawowych założeniach: zanikającej wrażliwości na wzrastające zyski czy straty oraz awersji do strat. Malejąca wrażliwość występuje wówczas, gdy wraz ze wzrostem odległości od punktu referencyjnego maleją zmiany krańcowe wartości wyników. Zatem dwukrotnie różniczkowalna funkcja wartości w zakresie zysków (dodatnich argumentów) ma drugą pochodną nieujemną, a w zakresie strat (ujemnych argumentów) drugą pochodną niedodatnią. Stąd też kształt funkcji wartości $v(x)$ przedstawiony na rys. 2 jest konsekwencją przyjętego założenia o malejącej wrażliwości. Stosując określenia teorii oczekiwanej użyteczności mówi się, że jednostka w obliczu perspektywy zysku wykazuje awersję do ryzyka, zaś w obliczu perspektywy strat skłonność do ryzyka.



Rys. 2. Funkcja wartości

Wypukłość funkcja, której wykres przedstawiono na rys. 2, obrazuje pożądane własności dotyczące malejącej wrażliwości, natomiast większa stromizna wykresu dla argumentów ujemnych ilustruje awersję do strat. Założenie dotyczące awersji do strat wyraża silniejsze uczucie żalu z perspektywy straty x jednostek niż zadowolenie związane z perspektywą zysku x jednostek. Inaczej mówiąc, emocje związane ze stratą x jednostek są silniejsze niż emocje związane z ich pozyskaniem – strata bardziej boli. Dla funkcji wartości $v(x)$

awersja do strat modelowana jest nierównością $v'(x) < v'(-x)$. Kanemann i Tverski nie przywiązują większej uwagi do postaci analitycznej funkcji wartości, a jedynie do jej kształtu. Zaproponowali funkcję wartości o następującej postaci

$$v_{K-T}(x) = \begin{cases} -\lambda(-x)^\alpha, & x < 0 \\ x^\alpha, & x \geq 0 \end{cases}$$

gdzie współczynnik λ wyraża „siłę” wrażliwości podejmującego decyzję na perspektywę poniesienia straty. Współczynnik λ jest niezależny od jednostek x , wpływa na współczynnik nachylenia wykresu funkcji wartości dla $x < 0$ i jest interpretowany jako charakterystyka stopnia awersji do strat. Powyższa funkcja przyjmowana jest jako podstawowa funkcja wartości. Empiryczna ocena parametru λ została oszacowana na poziomie 2,25.

Założenie znikającej wrażliwości również uzasadnia krzywiznę wykresu funkcji wag typu odwróconego S (rys. 1) tak, jak to miało miejsce w teorii RDEU. Zatem kształt wykresu funkcji wartości i wag jest konsekwencją założeń o zachowaniu jednostek w procesie wyboru i w tym sensie mówi się o psychologicznych podstawach tej teorii. Jednostka charakteryzująca się funkcją wartości $v(x)$ i subiektywną percepcją prawdopodobieństwa $g(p)$ podczas dokonywania w zbiorze dwuwynikowych wariantów decyzyjnych wyboru ujawnia swoje preferencje, które są zgodne z maksymalizacją funkcjonału $V(X_w) = g(p_1) v(x_1) + g(p_2) v(x_2)$.

Teoria perspektywy dzięki założeniu fazy edycji wyjaśniła efekt odbicia, ale była powszechnie krytykowana ze względu na możliwość naruszenia aksjomatu przechodniości, jak również braku zgodności z zasadą dominacji stochastycznej oraz ograniczenia metodologii do wyborów dwuwynikowych. Mimo tych wszystkich zarzutów, dała ona impuls do rozwoju ekonomii i finansów behawioralnych. Dalsze eksperymenty decyzyjne i badania w ramach tego nurtu pozwoliły na tyle poprawić metodologię, że jej rozszerzenie nazywane kumulacyjną teorią perspektywy (CTP) [11] stanowi kanon behawioralnej teorii decyzji.

Model teorii CPT postrzegany jest jako swoista kombinacja konstrukcji modelowej RDEU i TP. Teoria ta jest zgodna z aksjomatem malejącej wrażliwości i awersji do strat oraz nie ogranicza się tylko do analizy wariantów dwuwynikowych, jak to było w PT, lecz do wariantów losowych o dowolnej liczbie wyników rozpatrywanych względem punktu referencyjnego. Relatywizm wyników, malejącą wrażliwość i awersja do strat to trzy zasady o podłożu psychologicznym, które stanowią aksjomatykę teorii perspektyw, jak i kumulacyjnej teorii perspektywy. W koncepcji tych teorii jednostki nie mają

stałego stosunku do ryzyka, a zakładana w przez von Neumana i Morgensterna awersja do ryzyka nie jest cechą uniwersalną. Kahneman i Tverski za pomocą bardzo prostych i przekonujących sytuacji decyzyjnych pokazali, że ludzie zachowują się zgodnie z zasadą 4 zmian stosunku do ryzyka (fourfold pattern of risk attitudes) zgodnych z „zasadą odbicia”. Efekt ten polega na tym, że jednostki w obliczu małych prawdopodobieństw zysków wykazują skłonność do ryzyka, małych prawdopodobieństw strat awersję do ryzyka, średnich i dużych prawdopodobieństw zysków wykazują awersję do ryzyka oraz średnich i dużych prawdopodobieństw strat skłonność do ryzyka.

Ocena wariantów decyzyjnych w CPT jest sumą oceny wyników postrzeganych w kategoriach strat i zysków. Ogólnie funkcja oceny jest postaci

$$V(X) = V^+(X^+) + V^-(X^-)$$

gdzie X^+ reprezentuje zyski, a X^- straty.

Jeśli istnieje funkcja wartości $v(x)$ i funkcja wag $g(p)$ spełniająca wcześniej omawiane postulaty oraz π^- oznacza wagi decyzyjne zależne od kumulacyjnych funkcji rozkładu związanych z wystąpieniem strat oznaczone

$$\pi^- = g\left(\sum_{k=1}^i p_k\right) - g\left(\sum_{k=1}^{i-1} p_k\right)$$

π^+ dualne wagi decyzyjne zależne od dekumulacyjnych funkcji rozkładu związanych z wystąpieniem zysków

$$\pi^+ = g\left(\sum_{k=i}^n p_{n-k}\right) - g\left(\sum_{k=i+1}^n p_{n-k}\right)$$

to funkcja

$$V(X) = \pi^- v(x^-) + \pi^+ v(x^+)$$

jest zgodna z obserwowaną relacją preferencji spełniającą założenia teorii perspektywy. Zgodność aksjomatyki z funkcją oceny teorii CPT wykazali Chateauneuf i Wakker (1999). Jak już wcześniej pokazano, deformacja prawdopodobieństw poprzez funkcję wag określoną na wartościach dystrybucyjności i wykorzystanie w funkcji oceny wariantu odpowiednio skonstruowanych wag decyzyjnych zapewnia zgodność z warunkiem dominacji stochastycznej. Zatem rola funkcji wag jest bardzo ważna zarówno w teorii RDEU jak i PT i CPT. Szerokie studia dotyczące funkcji wag oraz oszacowania ich parametrów prowadzili Lattimore, Baker i Witte (1992), Kahneman i Tversky (1992), Prelec (1998), Gonzales i Wu (1999). W tabeli 1 zestawiono podstawowe funkcje wag analizowane ze względu na występujące w nich parametry.

Tabela 1

Funkcje wag

$\frac{p_i^\beta}{[p_i^\beta + (1 - p_i)^\beta]^{1/\beta}}$	Kahneman i Tverski (1992)	$0 < \beta < 1$ kształt odwróconego S β – zmiana punktu przegięcia
$\frac{\alpha p_i^\beta}{\alpha p_i^\beta + \sum_k p_k^\beta}$	Lattimore, Baker i Witte (1992)	β – zmiana punktu przegięcia, $\beta < 1$ kształt odwróconego S $\alpha < 1$ powoduje $g(\sum p_i) < 1$
$\exp[-(-\ln p_i)^\beta]$	Prelec (1998)	$0 < \beta < 1$ kształt odwróconego S, stały punkt przegięcia $p = e^{-1}$
$\frac{\alpha p_i^\beta}{\alpha p_i^\beta + (1 - p_i)^\beta}$	Gonzales i Wu (1999)	α kontroluje wartości funkcji, β wpływa na krzywiznę wykresu

Podsumowanie

Tradycyjnie w literaturze z zakresu teorii decyzji, ze względu na odmienne założenia o podejmującym decyzję, mówi się o podziale na nurt normatywny i deskryptywny. Jednak teorie takie nie są wzajemnie niezależne i nie można ich rozdzielać w sposób radykalny. Dobry model deskryptywny opisujący zachowanie decydentów w sytuacji wyboru wariantu ryzykownego, ze względu na swoją wartość predykcyjną, staje się modelem normatywnym. I odwrotnie, przyjmowany przez osoby jakiś model normatywny spełnia funkcje deskryptywne, a zatem pozwala przewidywać rzeczywiste zachowania jednostek. Istniejący podział na teorie normatywne i deskryptywne jest nieostry i proponuje się rozróżniać inne dwa podejścia: konwencjonalne i niekonwencjonalne. Podział taki uwzględnia indywidualną racjonalność w procesie wyboru, a więc to, co z punktu widzenia podejmującego decyzję stanowi o jego istocie.

Celem powstania konwencjonalnych teorii stanowiących rozszerzenie EU było uwzględnienie, obserwowanych w baniach empirycznych, odchyłeń i niezgodności z jej aksjomatyką bez konieczności eliminacji pożądanych własności tej teorii. W konwencjonalnym nurcie teorii decyzji racjonalność definiowana jest za pomocą zbioru formalnych aksjomatów, a zjawiska probabilistyczne są źródłem ryzyka. W podejściach tych jednostka ma określoną funkcję preferencji w zbiorze decyzji i powinna zachowywać się tak, jakby w swoich wyborach kierowała się maksymalizacją tej funkcji. Konwencjonalne postulaty racjonal-

ności von Neumanna i Morgensterna nie uwzględniają behawioralnych czynników w procesie wyboru, co powoduje, że obserwowane w eksperymentach zachowania przeczą zgodności z zasadą maksymalnej użyteczności.

Niekonwencjonalne teorie, w swoim podstawowym założeniu, powinny uwzględniać proces ludzkiego myślenia i zachowanie jednostki w sytuacji wyboru, zaś racjonalność nie może być postrzegana jako obiektywny atrybut aktu wyboru. Racjonalność nie jest definiowana w sposób matematyczny, ale modelowana na podstawie obserwacji zachowań jednostek w różnych sytuacjach decyzyjnych. Stąd też wzorzec indywidualnych zachowań warunkuje w zasadzie racjonalność wyboru. Decyzje podejmowane są w określonym czasie i otoczeniu, a decydenci różnią się wiedzą, doświadczeniem i w różny sposób ulegają emocjom. Teorie niekonwencjonalne mieszczą się w nurcie behawioralnym, ponieważ modele i zasady wyboru są formułowane na podstawie obserwacji rzeczywistych zachowań jednostek. Zasadnicza różnica polega na tym, że w teoriach konwencjonalnych ekonomiści mówią, że jednostki zachowują się tak, jakby wzorcowy model wyboru istniał, natomiast w teoriach behawioralnych modelować można tylko zachowanie ludzi w procesie podejmowania decyzji.

Do teorii niekonwencjonalnych można zaliczyć teorię perspektywy wraz z jej rozszerzeniami i modyfikacjami jak również teorię ograniczonej racjonalności [10].

Niezależnie od tego, czy racjonalność definiowana jest aksjomatycznie czy w sposób behawioralny, w większości teorii formuła zasady decyzyjnej ma strukturę sumy ważonych użyteczności, a to stanowi uogólnienie zasady maksymalnej użyteczności. Inaczej mówiąc, funkcjonal $V(X)$ wartościujący warianty decyzyjne jest funkcją oceny wyniku i wag decyzyjnych odpowiednio konstruowanych.

Literatura

1. Chateauneuf A., Walker P. (1999). An Axiomatization of Cumulative Prospect Theory for Decision Under Risk. *Journal of Risk and Uncertainty*, 18, s. 137-145.
2. Fishburn P.P. (1978). Stochastic Dominance Without Transitive Preferences. *Management Science*, 24, s. 1268-1277.
3. Friedmana M., Savage L. (1948). The Utility Analysis of Choices Involving Risk. *Journal of political Economy*, 56, s. 279-304.
4. Gonzales R., Wu.G. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology*, 38, s. 129-166.

5. Kahneman D., Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47, s. 263-291.
6. Lattimore P.K., Baker J.R., Witte A.D. (1992). The Influence of Probability on Risky Choice. *Journal of Economic Behavior Organization*, 17, s. 377-400.
7. Prelec D. (1998). The Probability Weighting Function, *Econometrica*, 66, s. 497-527.
8. Quiggin J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization*, 35, s. 160-169.
9. Starmer C. (2000). Developments In Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *Journal of Economic Literature*, 38, s. 332-382.
10. Todd P., Gigerenzer G. (2003). Bounding Rationality to the World. *Journal of Economic Psychology*, 24, 2, s. 143-165.
11. Tversky A., Kahneman D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5, p. 297-323.
12. Yaari M.E. (1987). The Dual Theory of Choice Under Risk. *Econometrica*, 55, s. 95-115.

PATTERN OF THE EXPECTED UTILITY MODEL AND THE DEVELOPMENT OF DECISION THEORY

Summary

Two main approaches dominate in the theory of risky choices. The well known standard models are consistent with the paradigm of the expected utility theory. All theories based on the formal axioms as rules of rational behavior and the utility preference function which is designed to permit observed violations of the independence axiom are called conventional theories. The models of risky choices which can not be reduced to, or expressed purely on terms of a single preference function are called nonconventional. Such theories based on the observed behavior of individuals engaged in risky choices. The nonconventional approach corresponds with the mental activities actually involved in making choices. This paper reviews the main approaches that have been adopted to analyze and describe decisions rules in the context of expected utility theory methodology.

Joanna Olbryś

Politechnika Białostocka

WIELOCZYNNIKOWE HYBRYDOWE MODELE MARKET-TIMING POLSKICH FUNDUSZY INWESTYCYJNYCH*

Wprowadzenie

Fama i French (1993) oraz Carhart (1997) zaproponowali nowe zmienne objaśniające SMB, HML oraz WML, nazwane portfelami naśladowczymi (mimicking portfolios), które mogą być wykorzystane jako dodatkowe czynniki rynkowe w analizie wrażliwości portfeli akcyjnych funduszy inwestycyjnych. Przykłady takiego zastosowania czynników Famy i Frencha oraz Carharta dotyczą wieloczynnikowych modeli wspomagających ocenę umiejętności menadżerów funduszy w zakresie stosowania techniki market-timing. Zmodyfikowane modele trójczynnikiowe H-M-FF oraz T-M-FF polskich OFI akcji zostały już skonstruowane [20]. Obecnie celem badań jest konstrukcja nowej zmiennej rynkowej *WML* (Winners-minus-Losers), reprezentującej czynnik momentum Carharta na polskim rynku oraz wprowadzenie tej zmiennej do hybrydowych, wieloczynnikowych modeli market-timing. Dodatkowo celem utworzenia modeli hybrydowych jest zastąpienie zmodyfikowanych modeli Treynora-Mazuya oraz Henrikssona-Mertona jednym modelem poprzez zastosowanie nowej zmiennej, tzw. wartości dodanej, będącej efektem perfekcyjnego stosowania techniki market-timing przez zarządzającego portfelem [7]. Badanie obejmuje grupę 15 OFI akcji w okresie styczeń 2003-grudzień 2010. Zastosowano metodę SUR (Seemingly Unrelated Regression) [25] do estymacji zaproponowanych hybrydowych modeli market-timing w badanej grupie funduszy.

* Praca naukowa finansowana ze środków na naukę w latach 2009-2011 jako projekt badawczy własny Nr N N113 173237.

1. Konstrukcja portfeli naśladowujących Famy i Frencha na polskim rynku

Fama i French obserwowali relację między wartością księgową kapitału własnego spółki BV (Book Value) a wartością rynkową jej akcji MV (Market Value) [5]. Wartość księgową w znacznej mierze opiera się na danych historycznych, czyli nie uwzględnia perspektyw firmy w tak znaczącym stopniu, jak bieżąca wartość rynkowa [8]. W przypadku każdej ze spółek Fama i French cyklicznie wyznaczali wskaźnik BV/MV, czyli stosunek wartości księgowej do wartości rynkowej. Spółki o niskiej wartości wskaźnika BV/MV były klasyfikowane jako spółki o potencjale wzrostu, natomiast te o wysokiej wartości BV/MV – jako spółki o potencjale wartości. Okazało się, że akcje spółek o potencjale wzrostu były bardziej ryzykowne i osiągały niższe stopy zwrotu, niż akcje spółek o potencjale wartości [8].

W kolejnym artykule Fama i French przedstawili trójczynnikiowy model równowagi cenowej akcji, w którym jako zmienne objaśniające zaproponowali: nadwyżkę rynkowej stopy zwrotu nad wolną od ryzyka stopą zwrotu, czynnik SMB (Small-minus-Big) – czynnik skonstruowany głównie na podstawie wartości rynkowej MV, tzw. size factor, oraz HML (High-minus-Low) – czynnik skonstruowany głównie na podstawie wartości wskaźnika BV/MV, tzw. book-to-market factor. Czynniki SMB oraz HML, uwzględniały zaobserwowaną wcześniej znaczącą wartość informacyjną wskaźnika BV/MV.

Portfele naśladowujące SMB oraz HML mogą być wykorzystane nie tylko jako dodatkowe czynniki rynkowe w analizie portfeli akcji, ale również w przypadku portfeli akcyjnych funduszy inwestycyjnych. Przykłady takiego zastosowania czynników Famy i Frencha dotyczą wieloczynnikowych modeli wspomagających ocenę umiejętności menedżerów funduszy w zakresie stosowania techniki market-timing [2]. Czynniki SMB i HML na polskim rynku zostały skonstruowane przez Olbryś [20], z wykorzystaniem procedury zaproponowanej przez Famę i Frencha [5]. Zaktualizowano bazę spółek. Usunięto z niej spółkę IGROUP z powodu ujemnej wartości księgowej na 1 akcję, na koniec 2009 roku. W bazie pozostało 60 spółek. Ceny zamknięcia uzupełniono korzystając z archiwum <http://bossa.pl>. Dane z raportów rocznych spółek z lat 2001-2009 (liczba akcji, wartość księgowa) pobrano ze strony www.bankier.pl (na podstawie serwisu Notoria). Wartość rynkową MV na 1 akcję stanowiła cena zamknięcia z ostatniego dnia roboczego czerwca danego roku. Wykonano

dodatkowe sortowanie spółek według wartości wskaźników MV oraz BV/MV 30 czerwca 2010 roku. Biorąc pod uwagę, że kolejne cykliczne sortowania wykonano odpowiednio w ostatnim dniu roboczym czerwca w latach 2003-2010, ze względu na dostępność danych z raportów rocznych za poprzedni rok, łącznie wykonano sortowanie 9 razy.

Spółki posortowane według wartości rynkowej MV były następnie dzielone na dwie grupy:

- spółki duże (B-Big), dla których wartość rynkowa akcji była nie mniejsza od mediany,
- spółki małe (S-Small), dla których wartość rynkowa akcji była mniejsza od mediany.

Spółki posortowane według wartości wskaźnika BV/MV były dzielone na trzy grupy:

- spółki, dla których wartość wskaźnika BV/MV była nie mniejsza od percentyla 70% (grupa H-High),
- spółki, dla których wartość wskaźnika BV/MV była mniejsza od percentyla 70%, ale nie mniejsza od percentyla 30% (grupa M-Medium),
- spółki, dla których wartość wskaźnika BV/MV była mniejsza od percentyla 30% (grupa L-Low).

W kolejnym kroku procedury odbywał się podział spółek na sześć rozłącznych grup: BH, BM, BL, SH, SM, SL. Do każdej z grup wchodziły te spółki, które spełniały oba warunki jednocześnie, czyli np. do grupy BH spółki duże, o wysokiej wartości wskaźnika BV/MV. Otrzymano w ten sposób po sześć, ważonych wartościami rynkowymi, portfeli pomocniczych w każdym roku. Pomiędzy datami kolejnych przetasowań skład portfeli pomocniczych pozostawał ten sam.

Ostatnim krokiem procedury było utworzenie portfeli naśladujących SMB i HML oraz wyznaczenie ich dziennych stóp zwrotu w okresie 2.01.2003-31.12.2010 według wzorów

$$R_{SMB} = \frac{1}{3} \cdot (R_{SH} + R_{SM} + R_{SL} - R_{BH} - R_{BM} - R_{BL}) \quad (1)$$

$$R_{HML} = \frac{1}{2} \cdot (R_{BH} + R_{SH} - R_{BL} - R_{SL}) \quad (2)$$

2. Konstrukcja czynnika momentum na polskim rynku

Strategie momentum polegają na możliwości osiągnięcia przez inwestorów ponadprzeciętnych zysków dzięki nabywaniu tych akcji, których ceny w ostatnim okresie znacząco zyskały na wartości (Winners), a unikaniu tych których ceny w największym stopniu zmniejszyły się (Losers) [3]. Jegadeesh i Titman w pracy z 1993 roku stwierdzili występowanie efektu momentum na amerykańskim rynku akcji w latach 1965-1989. W 1997 roku Carhart [4] zaproponował konstrukcję czteroczynnikowego modelu wyjaśniającego kształtowanie się stóp zwrotu z portfeli funduszy inwestycyjnych, w którym, oprócz czynników Fama i Frencha SMB i HML, uwzględnił efekt momentum, zidentyfikowany przez Jegadeesha i Titmana. Rouwenhorst [23] testował efekt momentum w 12 krajach europejskich* w latach 1980-1995 i uzyskał wyniki w większości zgodne z wynikami Jegadeesha i Titmana. Z kolei, analizując polski rynek, Buczek [3] stwierdził występowanie efektu momentum na Warszawskiej Giełdzie Papierów Wartościowych w latach 2001-2004.

W celu zbadania istotności wpływu zmiennej momentum oraz wprowadzenia tej zmiennej do hybrydowych, wieloczynnikowych modeli market-timing, dokonano konstrukcji nowej zmiennej objaśniającej WML (Winners-minus-Losers), reprezentującej czynnik momentum na polskim rynku. Czynniki WML został skonstruowany z wykorzystaniem procedury przedstawionej przez Carharta [4]. Propozycję oznaczenia czynnika jako WML zaczerpnięto z pracy Liew, Vassalou (2000).

W pierwszej kolejności dokonano selekcji spółek giełdowych, które uwzględniono w bazie danych na podstawie następujących kryteriów:

1) spółka była notowana na GPW w Warszawie co najmniej od 31 grudnia 2001 roku, aby dostępna była cena akcji w tym dniu;

2) dane dzienne (ceny zamknięcia) spółki miały możliwe do uzupełnienia luki; luki uzupełniono poprzez wykorzystanie średniej arytmetycznej sąsiednich notowań; niektóre ze spółek, np. Optimus, spełniały warunek 1, ale luki w danych były zbyt duże.

Do bazy danych utworzonej zgodnie z warunkami 1-2 weszło ostatecznie 71 spółek. W celu konstrukcji wartości indeksu WML obliczono 11-miesięczne, opóźnione o 1 miesiąc, stopy zwrotu. Obliczenia wykonano na koniec każdego miesiąca, od 31.12.2002 do 31.12.2010, wykorzystując dane dzienne. W następnej kolejności dokonano sortowania spółek według wartości 11-mie-

* Badane kraje to: Austria, Belgia, Dania, Francja, Niemcy, Włochy, Niderlandy, Norwegia, Hiszpania, Szwecja, Szwajcaria, Wielka Brytania [23].

sięcznej, opóźnionej o 1 miesiąc, stopy zwrotu, w ostatnim dniu roboczym każdego miesiąca. Łącznie wykonano sortowanie 96 razy. Spółki posortowane według wartości 11-miesięcznej, opóźnionej stopy zwrotu były dzielone na trzy grupy:

- spółki, dla których wartość 11-miesięcznej, opóźnionej stopy zwrotu była nie mniejsza od percentyla 70% (grupa W-Winners); w skład tej grupy weszły 22 spółki o najwyższych wartościach 11-miesięcznej, opóźnionej stopy zwrotu,
- spółki, dla których wartość 11-miesięcznej, opóźnionej stopy zwrotu była mniejsza od percentyla 70%, ale nie mniejsza od percentyla 30%; ta grupa spółek nie brała udziału w tworzeniu wartości indeksu WML,
- spółki, dla których wartość 11-miesięcznej, opóźnionej stopy zwrotu była mniejsza od percentyla 30% (grupa L-Losers); w skład tej grupy weszły 22 spółki o najniższych wartościach 11-miesięcznej, opóźnionej stopy zwrotu.

Pomiędzy datami kolejnych przetasowań skład portfeli pomocniczych W oraz L pozostawał nie zmieniony. Ostatnim krokiem procedury było utworzenie portfela WML jako różnicy średnich ważonych stóp zwrotu spółek z grup W i L oraz wyznaczenie dziennych stóp zwrotu z tego portfela w okresie 2.01.2003-31.12.2010 według wzoru

$$R_{WML} = \frac{1}{22} \cdot (\sum R_W - \sum R_L) \quad (3)$$

3. Trójczynnikowe modele market-timing

Trójczynnikowa modyfikacja klasycznego parametrycznego modelu market-timing Treynora-Mazuya, uwzględniająca czynniki rozpiętościowe SMB i HML Famy i Frencha, czyli model T-M-FF, ma postać [18]

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMB,t} + \delta_{2P} \cdot r_{HML,t} + \gamma_P \cdot (r_{M,t})^2 + \varepsilon_{P,t} \quad (4)$$

gdzie:

- | | |
|-----------------------------------|--|
| $r_{P,t} = R_{P,t} - R_{F,t}$ | jest nadwyżką zwykłej stopy zwrotu z portfela P nad wolną od ryzyka stopą zwrotu w okresie t , |
| $r_{M,t} = R_{M,t} - R_{F,t}$ | jest nadwyżką zwykłej stopy zwrotu z portfela rynkowego M nad wolną od ryzyka stopą zwrotu w okresie t , |
| $r_{SMB,t} = R_{SMB,t} - R_{F,t}$ | jest nadwyżką zwykłej stopy zwrotu z portfela naśladującego SMB nad wolną od ryzyka stopą zwrotu w okresie t , |

- $r_{HML,t} = R_{HML,t} - R_{F,t}$ jest nadwyżką zwykłej stopy zwrotu z portfela naśladowanego HML nad wolną od ryzyka stopą zwrotu w okresie t ,
- α_P jest miarą umiejętności zarządzającego portfelem P w zakresie selektywności aktywów – współczynnik alfa Jensena [12],
- β_P jest miarą ryzyka systematycznego portfela P ,
- γ_P jest miarą umiejętności zarządzającego portfelem P w zakresie stosowania techniki market-timing [9],
- δ_{1P} jest miarą wrażliwości stopy zwrotu z portfela P na zmiany stopy zwrotu portfela SMB,
- δ_{2P} jest miarą wrażliwości stopy zwrotu z portfela P na zmiany stopy zwrotu portfela HML,
- $\varepsilon_{P,t}$ jest składnikiem losowym, spełniającym następujące standardowe założenia modelu CAPM: $E(\varepsilon_{P,t}) = 0; E(\varepsilon_{P,t} | \varepsilon_{P,t-1}) = 0$.

Model H-M-FF, czyli trójczynnikowa modyfikacja klasycznego parametrycznego modelu market-timing Henrikssona-Mertona, ma postać [18]

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMB,t} + \delta_{2P} \cdot r_{HML,t} + \gamma_P \cdot y_{M,t} + \varepsilon_{P,t} \quad (5)$$

gdzie:

$$y_{M,t} = \max\{0, R_{F,t} - R_{M,t}\} = \max\{0, -r_{M,t}\}$$

Pozostałe oznaczenia jak we wzorze (4).

4. Parametryczny test G-I-I (Goetzman-Ingresoll-Ivkovič) – modyfikacja klasycznego modelu market-timing

Zmodyfikowane trójczynnikowe modele T-M-FF (4) i H-M-FF (5) mogą być szacowane zarówno na podstawie danych miesięcznych, jak i dziennych. Bollen i Busse [2] zwrócili uwagę na fakt, że większość wcześniejszych badań wspomagających analizę efektywności zarządzania portfelami funduszy inwestycyjnych na różnych rynkach opierała się na danych miesięcznych lub rocznych, natomiast analiza danych dziennych znacznie zmieniła wyniki badań. Na polskim rynku potwierdzają ten fakt wyniki Olbryś [22]. Goetzmann, Ingersoll i Ivković z kolei analizowali problem stosowania w modelach klasycznych danych miesięcznych w przypadku zarządzających, którzy codziennie (lub częściej niż raz w miesiącu) podejmują decyzje związane ze stosowaniem techniki market-timing. Zaproponowali odpowiednią modyfikację klasycznego modelu Henrikssona-Mertona – test G-I-I, polegającą na wprowadzeniu do modelu market-timing nowej zmiennej objaśniającej $P_{M,t}$ w miejsce zmiennej objaś-

niającej $y_{M,t} = \max\{0, -r_{M,t}\}$ (wzór (5)). Wartość dodana $P_{M,t}$ w miesiącu t , będąca efektem perfekcyjnego stosowania techniki market-timing przez zarządzającego portfelem, wyraża się wzorem [7]

$$P_{M,t} = \left[\left(\prod_{\tau=1}^N \max\{1 + R_{M,\tau}; 1 + R_{F,\tau}\} \right) - 1 \right] - R_{M,t} \quad (6)$$

gdzie:

N – liczba dni roboczych w miesiącu t , $\tau = 1, 2, \dots, N$.

Do oszacowania wartości zmiennej objaśniającej $P_{M,t}$ (6) potrzebne są dzienne oraz miesięczne stopy zwrotu z portfela rynkowego ($R_{M,t}$), jak również wartości dziennej wolnej od ryzyka stopy zwrotu ($R_{F,t}$).

5. Wieloczynnikowy hybrydowy model market-timing

Hybrydowy model czteroczynnikowy został skonstruowany przez Olbryś [21] jako modyfikacja modeli trójczynnikowych T-M-FF oraz H-M-FF, polegająca na wprowadzeniu nowej zmiennej objaśniającej $P_{M,t}$ (6), będącej wartością dodaną w miesiącu t , w miejsce zmiennych reprezentujących w danym modelu efekt perfekcyjnego stosowania strategii market-timing przez zarządzającego funduszem, czyli:

- kwadratu rynkowej zmiennej objaśniającej $(r_{M,t})^2$ w modelu T-M-FF (4),
- zmiennej $y_{M,t}$ w modelu H-M-FF (5).

W rezultacie otrzymano model postaci [21]

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMB,t} + \delta_{2P} \cdot r_{HML,t} + \gamma_P \cdot P_{M,t} + \varepsilon_{P,t} \quad (7)$$

gdzie: oznaczenia jak we wzorach (4), (5) oraz (6).

W cytowanej wyżej pracy przeprowadzono estymację i weryfikację modeli czteroczynnikowych, jak również interpretację uzyskanych wyników oceny dla 15 funduszy akcji na polskim rynku w okresie styczeń 2003-grudzień 2009. Dokonano również analizy porównawczej modeli trójczynnikowych oraz czteroczynnikowych. Zaproponowany czteroczynnikowy model hybrydowy okazał się przydatny do oceny efektywności zarządzania portfelami polskich funduszy akcji w badanym okresie. Model ten może zastąpić modele trójczynnikowe T-M-FF (4) oraz H-M-FF (5), ale wymaga zastosowania zarówno dziennych, jak i miesięcznych danych empirycznych [2]. Jednocześnie konieczność wykorzystania do konstrukcji zmiennych objaśniających danych o różnej częstotliwości można potraktować jako zaletę tego modelu.

Głównym celem niniejszej pracy jest rozszerzenie modelu market-timing do wersji pięcioczynnikowej, zawierającej jako dodatkową zmienną objaśniającą portfel naśladowujący WML (3), reprezentujący czynnik momentum na polskim rynku. Hybrydowy model pięcioczynnikowy przyjmie zatem postać

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMB,t} + \delta_{2P} \cdot r_{HML,t} + \delta_{3P} \cdot r_{WML,t} + \gamma_P \cdot P_{M,t} + \varepsilon_{P,t} \quad (8)$$

gdzie:

$r_{WML,t} = R_{WML,t} - R_{F,t}$ jest nadwyżką zwykłej stopy zwrotu z portfela naśladowującego WML nad wolną od ryzyka stopą zwrotu w okresie t ,

δ_{3P} jest miarą wrażliwości stopy zwrotu z portfela P na zmiany stopy zwrotu portfela WML.

Pozostałe oznaczenia jak we wzorach (4), (5), (6), (7).

6. Charakterystyka danych i wyniki empiryczne

Estymacja modelu pięcioczynnikowego (8) odbywa się z wykorzystaniem nadwyżek stóp zwrotu:

- nadwyżki miesięcznych stóp zwrotu z portfela danego funduszu oraz z portfela rynkowego nad wolną od ryzyka stopą zwrotu, czyli wartości zmiennych: objaśnianej $r_{P,t}$ oraz objaśniającej $r_{M,t}$; reprezentantem portfela rynkowego jest indeks WIG,
- nadwyżki średnich miesięcznych stóp zwrotu każdego z portfeli naśladowujących SMB, HML oraz WML nad wolną od ryzyka stopą zwrotu, czyli wartości zmiennych $r_{SMB,t}$, $r_{HML,t}$ i $r_{WML,t}$ odpowiednio,
- wartości zmiennej objaśniającej $P_{M,t}$ (6).

Wymienione powyżej stopy proste wyznaczane są z wykorzystaniem średniej miesięcznej rentowności bonów skarbowych 52-tygodniowych jako miesięcznej wolnej od ryzyka stopy zwrotu. Zbadano stacjonarność procesów $r_{M,t}$, $r_{SMB,t}$, $r_{HML,t}$, $r_{WML,t}$ oraz $P_{M,t}$ na podstawie rozszerzonego testu Dickeya-Fullera (ADF) i odrzucono we wszystkich przypadkach hipotezę o istnieniu pierwiastków jednostkowych.

Przeprowadzono estymację i weryfikację modeli pięcioczynnikowych dla 15 funduszy akcji na polskim rynku w okresie styczeń 2003-grudzień 2010, stosując metodę analizy danych panelowych, zaproponowaną przez Zellnera [25], a mianowicie estymację modelu pozornie niezwiązanych równań regresji SUR [15]. Metodę SUR można stosować, gdy dane panelowe charakteryzują się znaczną liczbą okresów oraz niewielką liczbą jednostek przekrojowych. W naszym badaniu: $T = 96$, $N = 15$. Podstawowe założenia modelu SUR są następujące [16]:

- 1) wartość oczekiwana składników losowych jest równa zero,
- 2) wariancja składników losowych jest stała w czasie dla danej jednostki przekrojowej, przy czym każda jednostka może mieć inną wariancję,
- 3) składniki losowe pochodzące z różnych jednostek przekrojowych, ale z tego samego okresu są skorelowane; jest to tzw. korelacja równoczesna (contemporaneous correlation),
- 4) składniki losowe pochodzące z różnych okresów nie są skorelowane, niezależnie od tego, czy dotyczą tej samej czy różnych jednostek przekrojowych.

W przypadku modeli grupy funduszy inwestycyjnych z tej samej klasy ryzyka mogą wystąpić efekty korelacji równoczesnej składników losowych (założenie 3) z powodu podobieństw składu portfeli funduszy [16]. Może być to wskazaniem do zastosowania metody SUR w celu uzyskania zadowalających wyników estymacji [19].

Tabela 1 przedstawia wyniki estymacji hybrydowych modeli market-timing w badanej grupie funduszy, w okresie styczeń 2003-grudzień 2010, uzyskane z wykorzystaniem pakietu Gretl 1.8.5. Parametry istotnie różniące się od zera są oznaczone odpowiednio: * istotność na poziomie 0,1; ** istotność na poziomie 0,05; *** istotność na poziomie 0,01 [13].

Obserwujemy wysoki stopień dopasowania modeli do danych empirycznych, mierzony skorygowanym współczynnikiem determinacji $\overline{R^2}$. Średnia wartość tego współczynnika wynosi 0,94. Dla 10 spośród 15 funduszy otrzymano dodatnie oraz istotne statystycznie wartości estymatora $\hat{\alpha}_p$, co potwierdza uzyskane wcześniej wyniki [21]. Cechą charakterystyczną modeli market-timing jest również istotność wpływu zmiennej niezależnej, reprezentującej portfel rynkowy, na stopy zwrotu z portfeli funduszy (dodatnie wartości parametru $\hat{\beta}_p$). Wartości estymatorów $\hat{\delta}_{1p}$, $\hat{\delta}_{2p}$ i $\hat{\delta}_{3p}$ interpretujemy odpowiednio jako miary wrażliwości stopy zwrotu z portfela funduszu na zmiany stóp zwrotu portfeli SMB, HML i WML, naśladujących rolę wybranych czynników. Czynnik SMB reprezentuje średnią miesięczną rozpiętość stóp zwrotu pomiędzy spółkami o niskiej i wysokiej wartości rynkowej MV. Stwierdzono wyraźną istotność statystyczną wpływu tego czynnika w przypadku 12 z 15 funduszy (estymatory parametru $\hat{\delta}_{1p}$). Z kolei wartości zmiennej HML mierzą średnią miesięczną rozpiętość stóp zwrotu pomiędzy spółkami o potencjale wartości (wysokie BV/MV) oraz spółkami o potencjale wzrostu (niskie BV/MV). Nie stwierdzono istotnego wpływu tego czynnika w przypadku portfeli funduszy akcji polskich (wartości parametru $\hat{\delta}_{2p}$), co jest zgodne z wynikami dotychczasowych badań [21].

Tabela 1

Hybrydowe modele market-timing (8) (dane z okresu styczeń 2003–grudzień 2010)

	Fundusz akcji	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\delta}_{1P}$	$\hat{\delta}_{2P}$	$\hat{\delta}_{3P}$	$\hat{\gamma}_P$	$\overline{R^2}$
1	Arka BZ WBK FIO Subfundusz Arka BZ WBK Akcji	0,025***	0,830***	1,490**	1,407**	-0,543	-0,127***	0,93
2	Aviva Investors FIO Subfundusz Aviva Polskich Akcji	0,021***	0,828***	1,691***	0,684	-0,410	-0,102***	0,96
3	BPH FIO Parasolowy BPH Subfundusz Akcji	0,009*	0,776***	0,935**	-0,003	0,216	-0,050*	0,95
4	Investor Top 25 Małych Spółek FIO (d. DWS Top 25)	0,058***	0,633***	7,817***	2,92***	0,554	-0,163***	0,89
5	Investor Akcji Dużych Spółek FIO (d. DWS)	0,001	0,791***	-0,578	-0,226	-0,381	-0,080***	0,94
6	Investor Akcji FIO (d. DWS Plus)	0,016***	0,724***	1,312***	0,275	0,021	-0,111***	0,94
7	ING Parasol FIO ING Subfundusz Akcji	0,001	0,855***	0,711*	-0,255	0,359	0,002	0,95
8	Legg Mason Akcji FIO	0,014**	0,770***	0,703	-0,353	1,193*	-0,059*	0,93
9	Millennium FIO Subfundusz Akcji	0,010*	0,726***	1,427***	0,011	-0,278	-0,075**	0,92
10	Novo FIO Subfundusz Novo Akcji	0,005	0,796***	1,094**	0,188	-0,959	-0,064*	0,93
11	Pioneer FIO Subfundusz Pioneer Akcji Polskich	-0,006	0,970***	0,367	0,476	-1,337**	-0,023	0,96
12	PKO Akcji – FIO	0,021***	0,683***	1,128**	0,165	-0,190	-0,198***	0,94
13	PZU FIO Parasolowy Subfundusz PZU Akcji Krakowiak	0,015***	0,727***	1,328***	0,734	-0,079	-0,095***	0,94
14	Skarbiec FIO Subfundusz Akcji Skarbiec – Akcja	0,013**	0,701***	1,288**	0,012	-0,751	-0,092**	0,92
15	UniFundusze FIO Subfundusz UniKorona Akcje	0,007	0,822***	0,888*	-0,046	0,392	-0,014	0,94
	Średnia	0,014	0,772	1,440	0,399	-0,024	-0,080	0,94

Wartości trzeciej zmiennej naśladowującej WML są miarą średniej miesięcznej rozpiętości stóp zwrotu spółek o najwyższych i najniższych wartościach 11-miesięcznej, opóźnionej stopy zwrotu. Wpływ tej zmiennej (reprezentującej czynnik momentum na polskim rynku) na wartości stóp zwrotu z portfeli badanych funduszy akcji polskich okazał się nieistotny statystycznie (tabela 1, estymatory $\hat{\delta}_{3P}$). Ostatnią ze zmiennych objaśniających, występujących w hybrydowym modelu (8) jest wartość dodana $P_{M,t}$. Analizując wzór (6) zauważamy, że parametr $\hat{\gamma}_P$ należy zinterpretować jako miarę umiejętności zarządzającego portfelem funduszu w zakresie osiągnięcia nadwyżek stóp zwrotu w stosunku do stopy zwrotu z portfela rynkowego. Analogicznie zatem, jak w przypadku modeli (4) i (5), ujemna oraz istotna statystycznie wartość estymatora tego parametru (która występuje w przypadku 12 z 15 badanych funduszy) może oznaczać negatywny wpływ stosowania techniki market-timing na wartość stopy zwrotu z portfela [18].

Podsumowanie

Głównym celem pracy była konstrukcja portfela naśladowującego WML, reprezentującego czynnik momentum na polskim rynku oraz analiza istotności wpływu tego czynnika jako zmiennej niezależnej w modelach market-timing grupy agresywnych funduszy inwestycyjnych. Otrzymano wynik negatywny, czyli brak istotnego wpływu czynnika WML na stopy zwrotu z portfeli funduszy. Wynik ten różni się od wniosków Carharta dla rynku amerykańskiego [4], zatem kolejnym etapem badań powinna być analiza odporności uzyskanych wyników estymacji na zmianę długości przedziału czasowego, np. poprzez analizę stabilności parametrów otrzymanych modeli ekonometrycznych.

Literatura

1. Baltagi B.H. (2005). *Econometric Analysis of Panel Data*. John Wiley & Sons, Ltd.
2. Bollen N.P.B., Busse J.A. (2001). On the Timing Ability of Mutual Fund Managers. *The Journal of Finance*, 56, 3.
3. Buczek B.S. (2005). *Efektywność informacyjna rynków akcji. Teoria a rzeczywistość*. SGH, Warszawa.
4. Carhart M.M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52, 1.
5. Fama E.F., French K.R. (1992). The Cross – Section of Expected Stock Returns. *The Journal of Finance*, 47.

6. Fama E.F., French K.R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33.
7. Goetzmann W.N., Ingersoll J. Jr., Ivković Z. (2000). Monthly Measurement of Daily Timers. *Journal of Financial and Quantitative Analysis*, 35, 3.
8. Haugen R.A. (1999). Nowa nauka o finansach. Przeciwno efektywności rynku. WIG-Press, Warszawa.
9. Henriksson R., Merton R. (1981). On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills. *Journal of Business*, 54, 4.
10. Henriksson R. (1984). Market Timing and Mutual Fund Performance: An Empirical Investigation. *Journal of Business* 57.
11. Jegadeesh N., Titman S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48, 1.
12. Jensen M. (1968). The Performance of Mutual Funds in the Period 1945-1964. *The Journal of Finance*, 23.
13. Kufel T. (2007). *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu Gretl*. PWN, Warszawa.
14. Liew J., Vassalou M. (2000). Can Book-to-Market, Size and Momentum be Risk Factors that Predict Economic Growth? *Journal of Financial Economics*, 57.
15. Maddala G.S. (2008). *Ekonometria*. PWN, Warszawa.
16. Marshall B.R., Young M. (2003). Liquidity and Stock Returns in Pure Order-Driven Markets: Evidence from the Australian Stock Market. *International Review of Financial Analysis*, 12.
17. Merton R. On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts. *Journal of Business*, 54, 3.
18. Olbryś J. (2010). Three-Factor Market-Timing Models with Fama and French's Spread Variables. *Operations Research and Decisions*, 2.
19. Olbryś J. (2010). Orthogonalized Factors in Market-Timing Models of Polish Equity Funds. *Metody Ilościowe w Badaniach Ekonomicznych*, 11, 1.
20. Olbryś J. (2010). Czynniki Famy i Frencha w wieloczynnikowych modelach market-timing polskich funduszy inwestycyjnych. *Zeszyty Naukowe. Uniwersytet Szczeciński, Szczecin*, 29.
21. Olbryś J. (2010). Ocena efektywności zarządzania portfelem funduszu inwestycyjnego z wykorzystaniem wybranych wieloczynnikowych modeli market-timing. *Optimum. Studia Ekonomiczne*, 4(48).
22. Olbryś J. (2008). Ocena umiejętności stosowania strategii market-timing przez zarządzających portfelami funduszy inwestycyjnych a częstotliwość danych. *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego*, 10.

23. Rouwenhorst K.G. (1998). International Momentum Strategies. *The Journal of Finance*, 53.
24. Treynor J., Mazuy K. (1966). Can Mutual Funds Outguess the Market? *Harvard Business Review*, 44.
25. Zellner A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of American Statistical Association*, 57.

MULTIFACTOR HYBRID MARKET-TIMING MODELS OF POLISH MUTUAL FUNDS

Summary

Fama and French (1993) and Carhart (1997) proposed new independent variables in the regressions of market-timing models. They formed the mimicking portfolios SMB, HML and WML, respectively. The portfolios SMB and HML have been introduced as explanatory variables into regressions of Polish equity mutual funds' portfolios excess returns in Olbryś (2010). The main goal of this paper is to present modified multifactor hybrid version of market-timing model with Carhart's momentum factor WML (Winners-minus-Losers) and the value added by perfect daily timing of fund assets. The market-timing and selectivity abilities of fund managers are evaluated for the period January 2003-December 2010. We compare the Seemingly Unrelated Regression (SUR) results of the models and investigate their statistical properties.

Aneta Peryga

Honorata Sosnowska

Szkoła Główna Handlowa w Warszawie

INTUICYJNE I DEDUKCYJNE ROZUMOWANIA W PODEJMOWANIU DECYZJI PRZY UWZGLĘDNIENIU BODŹCÓW FINANSOWYCH I WSPÓŁPRACY GRUPOWEJ

Wprowadzenie

Niniejsze opracowanie poświęcone jest badaniom nad tym, czy rozumowania intuicyjne przebiegają zgodnie z zasadami dedukcyjnymi logiki formalnej i rachunku prawdopodobieństwa. Badania przedstawione w tym artykule stanowią kontynuację badań Sosnowskiej opisanych w artykułach Sosnowskiej [12; 13; 14].

To jak ludzie probabilistycznie rozumują ma wpływ na podejmowane przez nich decyzje i zgodność lub nie rzeczywistych sytuacji z prognozami. W prognozach bowiem korzystamy z metod formalnych, zaś w intuicyjnych rozumowaniach przy podejmowaniu decyzji – niekoniecznie. Stąd waga badań dla teorii podejmowania decyzji i podstaw mikroekonomii nad tym, jaka jest zgodność pomiędzy rozumowaniami intuicyjnymi i formalnymi.

O tym, że rozumowania intuicyjne często nie są zgodne z regułami formalnymi wiadomo w psychologii poznawczej od dawna [11]. Zagadnieniem dotyczącymi implikacji zajmował się Wason (1966), zagadnieniami dotyczącymi koniunkcji – Tversky i Kahneman (1983). Wyniki te są obecnie umieszczane także w podręcznikach logiki matematycznej [5] po to, aby pokazać specyfikę rozumowań formalnych.

Najbardziej znane zagadnienie to przedstawiony przez Kahnemana i Tverskiego (1983) problem Lindy, gdy badani mają uszeregować, biorąc pod uwagę prawdopodobieństwo danej sytuacji, stwierdzenia dotyczące kobiety o imieniu Linda. O Lindzie dostają następującą informację: „Linda ma 31, jest singlem, wygadana i bardzo inteligentna. Ukończyła filozofię. Jako studentka

była głęboko zaangażowana w sprawy dyskryminacji i sprawiedliwości społecznej. Brała udział w ruchu antynuklearnym^{*}. Wśród stwierdzeń znajdują się dwa następujące:

- 1) Linda jest kasjerką w banku,
- 2) Linda jest kasjerką w banku i aktywną działaczką feministyczną.

Badani uznali, że stwierdzenie 2 jest bardziej prawdopodobne niż stwierdzenie 1. Wynik ten jest sprzeczny z prawem dotyczącym koniunkcji zdarzeń i stąd też używana w takich sytuacjach nazwa „błąd koniunkcji” („conjunction fallacy”). Różnym podejściom do problemu Lindy poświęcono dziesiątki artykułów i pojawiają się coraz to nowe analizy. W tym opracowaniu odwołamy się do tych, które łączą się z naszym podejściem.

Kolejny poddany badaniom spójnik logiczny to alternatywa (suma, dyzjunkcja). Badany rzadziej i później niż koniunkcja, ale w ostatnich czasach rośnie zainteresowanie rozumowaniami z dyzjunkcją. Za podstawową pracę przyjmuje się artykuł Bar-Hillel i Neter^{*} (1993) o błędzie dyzjunkcji („disjunction fallacy”). W artykule tym, będącym nawiązaniem do badań Tverskiego i Kahnemana, opisane są eksperymenty, których wyniki pokazują, że badani prawdopodobieństwo sumy zdarzeń określają jako mniejsze niż prawdopodobieństwo pojedynczego zdarzenia, czyli naruszone jest prawo dotyczące prawdopodobieństwa sumy zdarzeń, a tym samym alternatywy (dyzjunkcji) zdań. Bar Hillel i Neter badały bardzo specjalne układy zdarzeń, w których występował związek pomiędzy kategorią nadrzędną i podrzędną. Kolejni badacze osłabili te założenia. Costello badał prawdopodobieństwa zdarzeń, z którymi mamy do czynienia codziennie – elementów prognozy pogody.

Chciałoby się po prostu powiedzieć: „ludzie nie rozumują racjonalnie”, ale te błędy popełniają także ludzie o wykształconych rozumowaniach formalnych. Trwają badania nad wyjaśnieniem przyczyn błędów koniunkcji i dyzjunkcji. Brainerd i inni (2010) są zdania, że przyczyną jest działanie pamięci epizodycznej, zależnej od sposobu opisu. Costello twierdzi, że ludzie działają racjonalnie, tyle że prawa rachunku prawdopodobieństwa rozszerzają o element wariacyjny.

Prace Sosnowskiej [12; 13; 14] miały na celu zbadanie, czy studenci mający za sobą kurs logiki również popełniają błąd koniunkcji i dyzjunkcji. Odpowiedź jest twierdząca. Badano również, czy na występowanie tych błędów ma wpływ poziom wiedzy matematycznej badanych. W tym przypadku zaobserwowano pewną zależność, ale lepsze wykształcenie matematyczne nie wyklucza błędów koniunkcji i dyzjunkcji. Analizowano też wpływ typu zdarzeń (pozytywne czy negatywne) na występowanie tych błędów. Okazało, że błędy te licznie występują w przypadkach obu typów zdarzeń.

^{*} Bardziej dokładny opis dokonań naukowych Mayi Bar-Hillel można znaleźć w artykule Malawskiego i Tyszkii (2010)

Charnes i inni [4] badali, jaki wpływ na występowanie błędu koniunkcji w klasycznym problemie Lindy mają bodźce pieniężne wypłacane badanym za poprawne odpowiedzi i możliwość współpracy w parach lub trójkach. W opisywanym przez nich eksperymencie bodźce pieniężne i możliwość współpracy zmniejszały częstość występowania błędów. Tu przedstawimy wyniki eksperymentu, który miał sprawdzić, czy wyniki Charnesa i innych powtórzą się w warunkach polskich. Potwierdził się wpływ bodźców pieniężnych, nie ma potwierdzenia wpływu współpracy grupowej. Ponadto badano, czy zachodzą hipotezy Costello dotyczące jednoczesnego występowania błędu dyzjunkcji i koniunkcji, a także podwójnej koniunkcji i podwójnej dyzjunkcji.

1. Hipotezy badawcze

Mamy dwa zdarzenia A i B . Rozważamy część wspólna zdarzeń $A \cap B$ oraz sumę zdarzeń $A \cup B$. Badana osoba określa prawdopodobieństwo sumy i części wspólnej zdarzeń w zależności od prawdopodobieństwa zdarzeń A i B .

Mówimy, że badana osoba popełnia błąd koniunkcji, gdy stwierdza, że $P(A \cap B) > P(A)$ lub $P(A \cap B) > P(B)$. Błąd koniunkcji jest błędem podwójnej koniunkcji [6], gdy $P(A \cap B) > P(A)$ i $P(A \cap B) > P(B)$.

Mówimy, że badana osoba popełnia błąd dyzjunkcji, gdy stwierdza, że $P(A \cup B) < P(A)$ lub $P(A \cup B) < P(B)$. Błąd dyzjunkcji jest błędem podwójnej dyzjunkcji [6], gdy $P(A \cup B) < P(A)$ i $P(A \cup B) < P(B)$.

W niniejszym opracowaniu opiszemy eksperyment poddający weryfikacji następujące hipotezy badawcze.

H1. Bodźce pieniężne zmniejszają część badanych popełniających błąd koniunkcji i dyzjunkcji.

H2. Współpraca w parach lub w trójkach zmniejsza część badanych popełniających błąd koniunkcji i dyzjunkcji.

H3. Występowanie błędu koniunkcji (odpowiednio – podwójnej koniunkcji) jest dodatnio skorelowane z występowaniem błędu dyzjunkcji (odpowiednio – podwójnej dyzjunkcji).

Hipotezy H1 i H2 zostały sformułowane i potwierdzone przez Charnesa i innych [4] na studentach amerykańskich. Hipoteza H3 została zaś wyprowadzona matematycznie przy założeniu podejścia wariacyjnego do praw rachunku prawdopodobieństwa przez Costello. W ekonomii eksperymentalnej przyjęte jest, że eksperymenty powtarza się na różnych grupach badanych. W analizowanym przez nas eksperymencie badani byli polscy studenci, mający za sobą wykłady z elementów logiki.

2. Eksperyment

Eksperyment został przeprowadzony jesienią 2010 roku na grupie 206 studentów I roku studiów ekonomicznych w Szkole Głównej Handlowej w Warszawie. Eksperyment miał miejsce na wykładzie z matematyki, studenci w ramach tego wykładu poznali uprzednio elementy rachunku zdań i kwantyfikatorów. Studenci byli podzieleni na dwie, pracujące w osobnych salach, grupy: grupę, w której nie stosowano żadnych bodźców pieniężnych (101 osób) i grupę, której powiedziano, że pewne odpowiedzi są poprawne i za właściwe odpowiedzi badany dostanie wypłatę pieniężną (105 osób). Wypłata wynosiła 4 zł w przypadku jednej poprawnej odpowiedzi, 8 zł za dwie poprawne. Ustalając poziom wypłat kierowano się ceną piwa (małe, duże), co z grubsza odpowiada skali wypłat amerykańskich. Studenci każdej z grup byli podzieleni na 3 podgrupy – pracujących samodzielnie oraz mogących współpracować w dobranych według umieszczonych na kwestionariuszu numerów parach i trójkach. Dobór do podgrup opierał się na zasadzie sąsiedztwa w siedzeniu w trakcie eksperymentu. Współpracujące grupy mogły przedstawić wspólny wynik lub wyniki indywidualne. Liczby studentów w każdej podgrupie przedstawia tabela 1.

Tabela 1

Liczba osób biorących udział w eksperymencie

Typ pracy	Liczba osób w grupie eksperymentalnej		
	z bodźcami	bez bodźców	razem
Indywidualnie	30	30	60
Dwójki	30	30	60
Trójki	45	41	86
Razem	105	101	206

Każdy badany otrzymywał do wypełnienia kwestionariusz, w którym najpierw miał uszeregować na specjalnie przygotowanej skali 6 chorób (AIDS, depresja, ptasia grypa, rak, schizofrenia, złamanie kończyny) pod kątem swej subiektywnej oceny prawdopodobieństwa zachorowania na te choroby. Następnie na tej skali miał umieścić zdarzenia „zachoruję na AIDS i raka” oraz „zachoruję na AIDS lub raka”.

Niżej podajemy przykład skali z wypełnionego kwestionariusza. Im wyżej na skali tym większe prawdopodobieństwo zdarzenia. Zdarzenia wpisane są kursywą. W wierszach z numerami należało umieścić zdarzenie polegające na zachorowaniu na daną chorobę. W pozostałych wierszach zdarzenia „zachoruję na AIDS i raka” oraz „zachoruję na AIDS lub raka”.

Lista zdarzeń uporządkowanych według prawdopodobieństwa

-
- 1.....*Rak*.....
-
- 2.....*Depresja*.....
-*Rak lub AIDS*.....
- 3.....*Złamanie kończyny*.....
- 4.....*AIDS*.....
-
- 5.....*Schizofrenia*.....
-*Rak i AIDS*.....
- 6.....*Ptasia grypa*.....
-

W zależności od tego, czy badany należał do grupy pracującej z bodźcami pieniężnymi lub bez i od rodzaju możliwej współpracy między badanymi kwestionariusz był opatrzony jedną z sześciu poniższych dodatkowych informacji.

Dla grup pracujących bez bodźców pieniężnych:

1. Wypełniając kwestionariusz nie należy z nikim się konsultować.
2. Wykonując część II może Pan/i konsultować się z osobą, której kwestionariusz ma taki sam numer jak Pański. Powinna siedzieć obok. Państwa odpowiedzi mogą być różne.
3. Wykonując część II może Pan/i konsultować się z osobami, których kwestionariusze mają taki sam numer jak Pański. Powinny siedzieć obok. Państwa odpowiedzi mogą być różne.

Dla grup pracujących z bodźcami pieniężnymi:

4. Wypełniając kwestionariusz nie należy z nikim się konsultować.

W części II niektóre odpowiedzi są poprawne. Za każdą poprawną odpowiedź można otrzymać 4 zł maksymalnie 8 zł. Wyплаты będą po opracowaniu wyników eksperymentu. Proszę podać numer indeksu aby określić odbiorcę wypłaty.

5. Wykonując część II może Pan/i konsultować się z osobą, której kwestionariusz ma taki sam numer jak Pański. Powinna siedzieć obok. Państwa odpowiedzi mogą być różne.

W części II niektóre odpowiedzi są poprawne. Za każdą poprawną odpowiedź można otrzymać 4 zł maksymalnie 8 zł. Wyплаты będą po opracowaniu wyników eksperymentu. Proszę podać numer indeksu aby określić odbiorcę wypłaty.

6. Wykonując część II może Pan/i konsultować się z osobami, których kwestionariusze mają taki sam numer jak Pański. Powinny siedzieć obok. Państwa odpowiedzi mogą być różne.

W części II niektóre odpowiedzi są poprawne. Za każdą poprawną odpowiedź można otrzymać 4 zł maksymalnie 8 zł. Wyплаты będą po opracowaniu wyników eksperymentu. Proszę podać numer indeksu aby określić odbiorcę wypłaty.

Pełen kwestionariusz podajemy w aneksie na końcu opracowania.

Po opracowaniu wyników badania studenci otrzymali wypłatę pieniężną.

3. Wyniki

Zgodnie z polską terminologią dotyczącą rachunku prawdopodobieństwa będzie tu mowa o iloczynie zdarzeń (zamiast koniunkcji) i sumie zdarzeń (zamiast dyzjunkcji).

Najpierw sprawdzono, czy badani w ogóle rozróżniają iloczyn i sumę zdarzeń. Okazało się, że tak. Znakomita większość badanych tak uszeregowała zdarzenia, że prawdopodobieństwo sumy było większe niż prawdopodobieństwo iloczynu. Wyniki liczbowe i procentowe (w nawiasach) przedstawia tabela 2.

Tabela 2

Związek między ocenami prawdopodobieństwa iloczynu i sumy

Typ pracy	Udział udzielających odpowiedzi			
	prawdopodobieństwo iloczynu mniejsze niż prawdopodobieństwo sumy	prawdopodobieństwo iloczynu większe niż prawdopodobieństwo sumy	prawdopodobieństw iloczynu równe prawdopodobieństwu sumy	brak odpowiedzi dot. sumy lub iloczynu
1	2	3	4	5
Indywidualnie	52 (86,67%)	3 (5,00%)	3 (5,00%)	2 (3,33%)
w tym:				
z bodźcami	27 (90,00%)	–	2 (6,67%)	1 (3,33%)
bez bodźców	25 (83,33%)	3 (10,00%)	1 (3,33%)	1 (3,33%)

cd. tabeli 2

1	2	3	4	5
W dwójkach	53 (88,33%)	2 (3,33%)	–	5 (8,33%)
w tym:				
z bodźcami	29 (96,67%)	–	–	1 (3,33%)
bez bodźców	24 (80,00%)	2 (6,67%)	–	4 (13,33%)
W trójkach	70 (81,40%)	3 (3,49%)	–	13 (15,12%)
w tym:				
z bodźcami	40 (88,89%)	1 (2,22%)	–	4 (8,89%)
bez bodźców	29 (70,73%)	2 (4,88%)	0	10 (24,39%)
Razem	174 (84,47%)	8 (3,88%)	3 (1,46%)	21 (10,19%)
w tym:				
z bodźcami	96 (91,43%)	1 (0,95%)	2 (1,90%)	6 (4,76%)
bez bodźców	78 (77,23%)	7 (6,93%)	1 (0,99%)	15 (14,85%)

Wpływ bodźców pieniężnych przedstawia się następująco. Ponad 90% badanych w przypadku bodźców pieniężnych i nieco ponad 2/3 w sytuacji braku tych bodźców określało dobrze prawdopodobieństwo iloczynu jako mniejsze od prawdopodobieństw obu zdarzeń. Gorsze wyniki dotyczą sumy zdarzeń (co jest sprzeczne z obserwacjami Bar-Hillel (1993)). W tym wypadku jest trochę ponad 1/3 poprawnych odpowiedzi w przypadku braku bodźców pieniężnych i nieco ponad 2/3, gdy te bodźce są stosowane. Wyniki liczbowe i procentowe (w nawiasach) podaje tabela 3.

Tabela 3

Wpływ bodźców pieniężnych na poprawność odpowiedzi

Typ pracy	Prawdopodobieństwo iloczynu:			Prawdopodobieństwo sumy:		
	większe niżprawdopodobieństwa zdarzeń	pomiędzyprawdopodobieństwami zdarzeń	mniejsze niżprawdopodobieństwa zdarzeń	większe niżprawdopodobieństwa zdarzeń	pomiędzyprawdopodobieństwami zdarzeń	mniejsze niżprawdopodobieństwa zdarzeń
Razem	10 (4,85%)	9 (4,37%)	169 (82,04%)	112 (54,37%)	61 (29,61%)	15 (7,28%)
w tym						
z bodźcami	6 (5,71%)	–	95 (90,48%)	74 (70,48%)	22 (20,95%)	4 (3,81%)
bez bodźców	4 (3,96%)	9 (8,91%)	74 (73,27%)	38 (37,62%)	39 (38,61%)	11 (10,89%)

Zatem można wyraźnie stwierdzić, że stosowanie bodźców pieniężnych zwiększa istotnie część poprawnych odpowiedzi, co potwierdzają statystyki testów niezależności (dla iloczynu $\chi^2 = 9,212$; $p = 0,002$; dla sumy zdarzeń $\chi^2 = 21,091$; $p < 0,001$) oraz sugerowane przez współczynniki korelacji Phi istotne, choć niewielkiej siły korelacje (o wartościach równych odpowiednio $\Phi = 0,224$ i $\Phi = 0,33$). W ten sposób hipoteza H1 została zweryfikowana pozytywnie zarówno w przypadku koniunkcji, jak i dyzjunkcji. Frakcje prawidłowych odpowiedzi w przypadku sumy i iloczynu różnią się istotnie, o czym świadczą wyniki testów proporcji (odpowiednio $Z = 4,732$ dla sumy oraz $Z = 3,217$ dla iloczynu zdarzeń). Jednocześnie błąd koniunkcji popełnia znacznie mniej badanych, co znajduje odzwierciedlenie w niskim współczynniku korelacji (mierzonym statystyką Phi) pomiędzy zmiennymi opisującymi wystąpienie pojedynczego błędu koniunkcji a dyzjunkcji ($\Phi = 0,069$; $p = 0,319$, korelacja nieistotna). Widoczna jest też nieistotność statystycznej analogicznych współczynników dla danych podgrup. Nie znajduje zatem potwierdzenia hipoteza H3 o tym, że występowanie błędów koniunkcji i dyzjunkcji jest dodatnio skorelowane. W przypadku błędu podwójnej koniunkcji i dyzjunkcji może co prawda istnieć słaba zależność ($\Phi = 0,197$; $p = 0,05$), jednak brak istotności dla poszczególnych klas odpowiedzi oraz niewielka ilość takich błędów w próbie pozwala przypuszczać, iż sugerowana zależność jest wynikiem błędu interpretacji (systematycznego błędu treści). Tym samym, ponieważ hipoteza H3 jest oparta na matematycznym wyprowadzeniu równoważności błędów koniunkcji i dyzjunkcji w przypadku zastosowania podejścia wariacyjnego do praw rachunku prawdopodobieństwa [6] należy sądzić, że badani podejścia wariacyjnego nie stosują.

Jeśli chodzi o wpływ współpracy na wyniki eksperymentu to rezultaty są zawarte w tabeli 4.

Tabela 4

Wpływ współpracy na poprawność odpowiedzi

Typ pracy	Prawdopodobieństwo iloczynu:			Prawdopodobieństwo sumy:		
	większe niż prawdopodobieństwa zdarzeń	pomiędzy prawdopodobieństwami zdarzeń	mniejsze niż prawdopodobieństwa zdarzeń	większe niż prawdopodobieństwa zdarzeń	pomiędzy prawdopodobieństwami zdarzeń	mniejsze niż prawdopodobieństwa zdarzeń
1	2	3	4	5	6	7
Razem	10 (4,85%)	9 (4,37%)	169 (82,04%)	112 (54,37%)	61 (29,61%)	15 (7,28%)
w tym indywidualnie	5 (8,33%)	2 (3,33%)	51 (85%)	36 (60%)	21 (35%)	2 (3,33%)

cd. tabeli 4

1	2	3	4	5	6	7
dwójki	2 (3,33%)	1 (1,67%)	51 (85%)	30 (50%)	19 (31,67%)	7 (11,67%)
trójki	3 (3,49%)	6 (6,98%)	67 (77,91%)	46 (53,49%)	21 (24,42%)	6 (6,98%)

Najlepsze wyniki uzyskano dla pracy indywidualnej, potem dla dwójek, a najgorsze dla trójek. Jest to sprzeczne z wynikami Charnesa i innych [3; 4], gdzie współpraca, zwłaszcza w trójkach, poprawiała wyniki. Brak jest jednak statystycznie istotnej korelacji pracy grupowej z wynikami. Nie można też odrzucić hipotez o niezależności zmiennych. Hipoteza H2 nie znalazła zatem potwierdzenia, co można tłumaczyć faktem sygnalizowanym przez badaczy, że umiejętność współpracy nie jest cechą polskich studentów. Bardziej skomplikowanie wygląda to, gdy uwzględnimy podział na sytuacje z bodźcami pieniężnymi i bez (tabele 5 i 6).

Tabela 5

Wpływ współpracy na poprawność odpowiedzi
przy braku bodźców pieniężnych

Typ pracy	Prawdopodobieństwo iloczynu:			Prawdopodobieństwo sumy:		
	większe niżprawdopodobieństwa zdarzeń	pomiędzyprawdopodobieństwami zdarzeń	mniejsze niżprawdopodobieństwa zdarzeń	większe niżprawdopodobieństwa zdarzeń	pomiędzyprawdopodobieństwami zdarzeń	mniejsze niżprawdopodobieństwa zdarzeń
Razem	4 (3,96%)	9 (8,91%)	74 (73,27%)	38 (37,62%)	39 (38,61%)	11 (10,89%)
w tym						
indywidualnie	3 (10%)	2 (6,67%)	24 (80%)	16 (53,33%)	11 (6,67%)	2 (3,33%)
dwójki	1 (3,33%)	1 (3,33%)	23 (76,67%)	8 (26,67%)	13 (20%)	6 (13,33%)
trójki	—	6 (14,63%)	27 (65,85%)	14 (34,15%)	15 (7,32%)	3 (19,51%)

Tabela 6

Wpływ współpracy na poprawność odpowiedzi
przy bodźcach pieniężnych

Typ pracy	Prawdopodobieństwo iloczynu:			Prawdopodobieństwo sumy:		
	większe niż prawdo- podobień- stwa zdarzeń	pomiędzy prawdo- podobień- stwami zdarzeń	mniejsze niż prawdo- podobień- stwa zdarzeń	większe niż prawdo- podobień- stwa zdarzeń	pomiędzy prawdo- podobień- stwami zdarzeń	mniejsze niż prawdo- podobień- stwa zdarzeń
Razem	6 (5,71%)	–	95 (90,48%)	74 (70,48%)	22 (20,95%)	4 (3,81%)
w tym						
indywidualnie	2 (6,67%)	–	27 (90%)	20 (66,67%)	10 (33,33%)	–
dwójki	1 (3,33%)	–	28 (93,33%)	22 (73,33%)	6 (20%)	1 (3,33%)
trójki	3 (6,67%)	–	40 (88,89%)	32 (71,11%)	6 (13,33%)	3 (6,67%)

Dla przypadku bez bodźców pieniężnych najczęściej poprawnych odpowiedzi mamy indywidualnie, potem w dwójkach, a na końcu w trójkach dla iloczynu oraz indywidualnie, w trójkach i dwójkach dla sumy. Dla przypadku z bodźcami pieniężnymi iloczyn wypadł najlepiej w dwójkach, potem indywidualnie, a na końcu w trójkach, zaś suma najlepiej w dwójkach, potem w trójkach, a na końcu w dwójkach. Brak ścisłej zależności potwierdza się w wynikach testów χ^2 oraz braku statystycznej istotności współczynników korelacji V-Cramera. Zatem nie ma reguły w jaki sposób bodźce pieniężne wpływają na to, czy współpraca zwiększa część poprawnych odpowiedzi.

Podsumowanie

Otrzymane wyniki dotyczące hipotez H1 i H2 wpisują się w koncepcję dwu systemów pracy mózgu, systemu 1 i systemu 2 [7; 8]. W systemie 1 rozumiemy intuicyjnie, stosujemy różne uproszczone metody rozumowań. W systemie 2 nasze myślenie staje się bardziej analityczne i dokładne. Błędy dyzjunkcji i koniunkcji popełniamy w systemie 1, gdy natomiast w grę wchodzi motywacja (w naszym badaniu bodźce pieniężne), to przestawiamy się na system 2. Przystawienie się na system 2 powoduje też konieczność uzasadniania swego zdania. Można przypuszczać, że w eksperymencie Charnesa i innych

studenci amerykańscy uzasadniali swoje wybory, natomiast w opisywanym tutaj eksperymencie polskim nie dało się obserwatorom zauważać specjalnych dyskusji i być może właśnie ich brak powoduje, że możliwość współpracy nie ma wpływu na zwiększenie części poprawnych wyników.

Jeśli zaś chodzi o hipotezę H3, to fakt, że wyniki różnią się od otrzymanych przez Costello może być spowodowany różnicami w przedmiocie badania. W eksperymencie Costello badani określali prawdopodobieństwo czynników meteorologicznych składających się na prognozę pogody w Irlandii (np. słonecznie i wietrznie). Można przyjąć, że według myślenia w systemie 1 poszczególnych badanych większość z tych czynników jest ze sobą jakoś skorelowana i stąd, a nie z podejścia wariacyjnego do rachunku prawdopodobieństwa, uzyskano korelacje pomiędzy występowaniem błędu dyzjunkcji i koniunkcji. W naszym badaniu alternatywy na ogół nie są uważane przez badanych za skorelowane i nie uzyskaliśmy korelacji między błędem koniunkcji i dyzjunkcji.

ANEKS

KWESTIONARIUSZ

Bierze Pani/Pan udział w badaniu dotyczącym prawdopodobieństwa porządkowego.

Badanie składa się z dwu części.

CZĘŚĆ I

Następujące choroby trzeba umieścić w wierszach oznaczonych numerami od 1 do 6 w zależności od prawdopodobieństwa zachorowania na tę chorobę, zaczynając od tej, na którą prawdopodobieństwo zachorowania jest według Pani/Pana największe, kolejno te o coraz mniejszym prawdopodobieństwie, a kończąc na tej, na którą prawdopodobieństwo zachorowania jest najmniejsze.

CHOROBY

AIDS, depresja, ptasia grypa, rak, schizofrenia, złamanie kończyny

Lista zdarzeń uporządkowanych według prawdopodobieństwa

.....

1.....

.....

2.....

.....

- 3.....
- 4.....
- 5.....
- 6.....

CZEŚĆ II

W „Liście zdarzeń uporządkowanych wg prawdopodobieństwa” znajdują się wiersze, które nie są oznaczone liczbami. Prawdopodobieństwo zdarzenia umieszczonego w takim wierszu mieści się pomiędzy prawdopodobieństwami zdarzeń w sąsiednich wierszach oznaczonych liczbami. Należy w nieoznaczonych liczbami wierszach umieścić poniższe zdarzenia w zależności od tego, jak Pan/i określa prawdopodobieństwo jego wystąpienia:

- a) zachoruję na AIDS i raka,
- b) zachoruję na AIDS lub raka.

Literatura

1. Bar-Hiller M., Neter E. (1993). How Alike Is It Versus How Likely Is It: A Disjunction Fallacy in Probability Judgments. *Journal of Personality and Social Psychology*, 65, s. 1119-1131.
2. Brainerd C.J., Reyna V.F., C. Aydin. (2010). Remembering in Contradictory Minds: Disjunction Fallacies in Episodic Memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36, s. 711-735.
3. Charnes G.B., Lewin D., Karni E. (2008). On the Conjunction Fallacy in Probability Judgment: New Experimental Evidence. Departmental Working Paper, Department of Economics, University of California, Santa Barbara.
4. Charnes G.B., Lewin D., Karni E. (2010). On the Conjunction Fallacy in Probability Judgment: New Experimental Evidence Regarding Linda. *Games and Economic Behavior*, 68, s. 551-556.
5. Chiswell I., Hodges W. (2007). *Mathematical Logic*. Oxford University Press, Oxford.
6. Costello F.J. (2009). Fallacies in Probability Judgments for Conjunction and Disjunction of Everyday Events. *Journal of Behavioral Decision Making*, 22, s. 235-251.
7. Kahneman D. (2003). Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice. W: T. Frangsmgr: Les Prix Nobel: The Nobel Prizes 2002. s. 449-489.

8. Kahneman D. (2003). Maps of Bounded Rationality. *Psychology for Behavioral Economics*. *American Economic Review*, 93, s. 1449-1475.
9. Kahneman D., Tversky A. (1983). Extensional Versus Intuitive Reasoning. The Conjunction Fallacy in Probability Judgement. *Psychological Review*, 90, s. 293-315.
10. Malawski M., Tyszka T. (2010). Maya Bar-Hillel. *Decyzje*, 14, s. 117-124.
11. Nęcka E., Orzechowski J., Szymura B. (2006). *Psychologia poznawcza*. PWN, Warszawa.
12. Sosnowska H. (2008). Ryzyko porządkowe a rachunek zdarzeń. W: *Modelowanie preferencji a ryzyko '08*. Red. T. Trzaskalik. AE, Katowice, s. 91-100.
13. Sosnowska H. (2008). Prawdopodobieństwo subiektywne a reguły rachunku prawdopodobieństwa. W: *Metody ilościowe w ekonomii*. Wydawnictwo Wyższej Szkoły Bankowej, Poznań, s. 2-22.
14. Sosnowska H. (2010). Reguły formalne ryzyka porządkowego w przypadku zdarzeń pozytywnych i negatywnych. W: *Modelowanie preferencji a ryzyko '10*. Red. T. Trzaskalik. AE, Katowice, s. 301-314.
15. Wason P.C. (1966). Reasoning. In: *New Horizons in Psychology*. Red. B.M. Foss. Harmondsworth, UK, Penguin.

CONJUNCTION AND DISJUNCTION FALLACIES IN CASE OF MONEY INCENTIVES AND COOPERATION

Summary

This paper presents the results of experiments continuing previous Sosnowska's research. Connections between formal and intuitive reasoning are studied, especially conjunction and disjunction fallacies. Previous research showed that formal and intuitive reasoning use different reasoning rules. American research (Charnes et alii) showed that money incentives and possibility of cooperation give better convergence between intuitive and formal reasoning. We use this method in our experiment conducted in Poland. We obtained a positive influence of money incentives and lack of an influence of cooperation.

Marcin Salamaga

Uniwersytet Ekonomiczny w Krakowie

ZASTOSOWANIE ANALIZY CONJOINT DO BADANIA PREFERENCJI MUZYCZNYCH

Wprowadzenie

Muzyka stanowi jedną z najwznioślejszych form artystycznych stworzonych przez człowieka. IX Symfonia Beethovena, Wielka Msza h-moll Bacha czy Requiem Mozarta należą do najwybitniejszych osiągnięć geniuszu ludzkiego i są stawiane w jednym rzędzie z największymi dziełami Michała Anioła, Fidasza czy Szekspira. Słuchanie muzyki wyzwala nieprzebrane pokłady emocji, wzniosłych uczuć i wewnętrznych przeżyć, podobnie jak kontemplowanie arcydzieła malarskiego, rzeźbiarskiego czy literackiego. Mimo to muzyka poważna uchodzi za sztukę elitarną, dostępną raczej dla wąskiego grona odbiorców. Wśród przyczyn tego zjawiska można wymienić niedostateczną promocję muzyki poważnej w środkach masowego przekazu jak i to, że sama muzyka wymaga od słuchacza pewnego wysiłku intelektualnego celem poznania jej prawdziwej wartości. Jest ona nie tylko źródłem głębokich doznań estetycznych, ale także uwarżliwia słuchacza i jak napisał znany publicysta muzyczny Jerzy Waldorff „muzyka łągodzi obyczaje” [12]. W dobie powszechnej komercjalizacji również muzyka poważna do pewnego stopnia stała się „towarem” podlegającym ekonomicznym prawom rynku. Uczestnikami rynku muzycznego są między innymi firmy fonograficzne, studia nagrań, środki masowego przekazu, organizatorzy imprez muzycznych i oczywiście słuchacze, melomani reprezentujący popytową stronę rynku. Z tego też powodu istotne wydaje się właściwe rozpoznanie gustów i preferencji muzycznych melomanów, aby trafić do nich z optymalnie dobranym „produktem muzycznym”, którego nośnikiem może być płyta kompaktowa, program muzyczny, jak i koncert. Empiryczne badania upodobań muzycznych są od dawna podejmowane w literaturze psychologicznej, socjologicznej, rzadziej marketingowej. Wymienić można w tym miejscu prace A. Białkowskiego i M. Grusiewicza [2], J. Ginocchio [3], D.J. Hargreavesa i A.C. Northa [7], S. Hallama, I. Cossa i M. Thauta [6], A.J. Lonsdale i A.C. Northa [9] czy B. Kamińskiej [8]. Niniejsze opracowanie również wpisuje się w nurt badań nad preferencjami

i gustami muzyczni osób słuchających muzyki. Podstawowym narzędziem badawczym wykorzystanym w tej pracy jest analiza conjoint, którą przeprowadzono na podstawie wyników badania ankietowego wśród użytkowników serwisu internetowego YouTube*. Zastosowanie analizy conjoint umożliwiło oszacowanie użyteczności częściowych respondentów, dzięki którym obliczono poziom ważności zmiennych przedstawiających gust, estetykę muzyczną i kulturę muzyczną badanej grupy melomanów. Ponadto oszacowane użyteczności częściowe, pozwoliły na segmentację melomanów według zbliżonych preferencji wyboru. Wykonanie analizy conjoint w każdym z wyróżnionych segmentów dało możliwość szczegółowego opisu preferencji muzycznych respondentów z uwzględnieniem ważności czynników, które współtworzą te preferencje. Otrzymane wyniki badań odnoszą się do specyficznej i stosunkowo wąskiej grupy melomanów. Mają one więc charakter ilustracyjny i mogą być użyteczne w prowadzeniu szerszych badań służących dostosowaniu oferty wydawnictw muzycznych, wydawnictw płytowych, repertuaru sal koncertowych czy programów muzycznych w radiu i telewizji do oczekiwań melomanów.

1. Metodologia badania preferencji muzycznych

Podstawową metodą badawczą wykorzystaną w niniejszej pracy do porównania zainteresowań muzycznych użytkowników portalu YouTube jest analiza conjoint. Procedura conjoint jest metodą klasyfikacji i analizy danych służących pomiaru preferencji ankietowanych osób [10]. Respondenci oceniają zbiór profilów opisanych wybranymi zmiennymi objaśniającymi (atrybutami), co umożliwia pozyskanie wiedzy o całkowitych preferencjach respondentów względem wyróżnionych profilów. W kolejnym etapie analizy conjoint obliczane są użyteczności częściowe poziomów atrybutów na podstawie dekompozycji użyteczności całkowitych przeprowadzonej na bazie wyników ocen preferencji [4]. W artykule użyteczności respondentów zostały obliczone na podstawie ocen parametrów następującej funkcji regresji [1]

$$\hat{Y}_k = b_{0k} + \sum_{p=1}^n b_{pk} X_{pk} + e \quad (1)$$

gdzie:

Y_k – zmienna wyrażająca ocenę preferencji k -tego respondenta,

$b_{0k}, b_{1k}, b_{2k}, \dots, b_{nk}$ – parametry równania regresji,

$X_1, X_2, \dots, X_n$ – zmienne sztuczne reprezentujące poziomy atrybutów,

e – składnik losowy modelu.

* <http://www.youtube.com/>

Respondenci porządkowali poszczególne profile przez nadanie im rang, więc wartości zmiennej Y są mierzone na skali porządkowej. Natomiast zmienne objaśniające są z reguły zmiennymi jakościowymi o dwóch bądź o trzech kategoriach, dlatego w modelu (1) wprowadzono zmienne sztuczne $X_1 - X_2$ wyrażające wpływ każdego poziomu zmiennej na ocenę przypisaną profilom przez respondenta. Konstrukcję zmiennych sztucznych z użyciem kodowania quasi-eksperymentalnego oraz sposób obliczenia użyteczności częściowych respondentów przedstawiono w tabelach 1 i 2 [1].

Tabela 1

Quasi-eksperymentalne kodowanie zmiennej objaśniającej
o dwóch stanach

Zmienna objaśniająca	Zmienna sztuczna X_p	Użyteczności częstkowe
Poziom I	1	$U_{j1}^k = b_{ks}$
Poziom II	-1	$U_{j2}^k = -b_{ks}$

Źródło: [10].

Tabela 2

Quasi-eksperymentalne kodowanie zmiennej objaśniającej
o trzech stanach

Zmienna objaśniająca	Zmienna sztuczna X_p	Zmienna sztuczna X_q	Użyteczności częstkowe
Poziom I	1	0	$U_{j1}^k = b_{pk}$
Poziom II	0	1	$U_{j2}^k = b_{qk}$
Poziom III	-1	-1	$U_{j3}^k = -(b_{pk} + b_{qk})$

Źródło: [10].

Częstkowe użyteczności wykorzystano następnie do obliczenia całkowitych użyteczności i -tego profilu i k -tego respondenta zgodnie z następującym wzorem [10]

$$U_{ik} = \sum_{j=1}^m U_{jl_j}^k + b_{0k} \quad (2)$$

gdzie:

$U_{jl_j}^k$ – użyteczność cząstkowa l -tego poziomu j -tej zmiennej i -tego profilu dla k -tego respondenta ($k = 1, 2, \dots, K$),

l_j^i – numer poziomu dla j -tej zmiennej i i -tego profilu,

$i = 1, \dots, n$ – numer profilu,

$j = 1, \dots, m$ – numer zmiennej,

b_{0k} – wyraz wolny modelu (1).

W celu obliczenia przeciętnej całkowitej użyteczności k -tego profilu wykorzystano formułę [10]

$$U_i = \frac{1}{K} \sum_{k=1}^K U_{ik} \quad (3)$$

Ponadto, cząstkowe użyteczności umożliwiły obliczenie relatywnej ważności każdej ze zmiennych objaśniających dla k -tego respondenta według następującego wzoru [10]

$$W_j^k = \frac{\max_{l_j} \{U_{jl_j}^k\} - \min_{l_j} \{U_{jl_j}^k\}}{\sum_{j=1}^m \left(\max_{l_j} \{U_{jl_j}^k\} - \min_{l_j} \{U_{jl_j}^k\} \right)} \cdot 100\% \quad (4)$$

Natomiast do wyznaczenia średniej ważności j -tej zmiennej zastosowano formułę [10]

$$W_j = \frac{1}{K} \sum_{k=1}^K W_j^k \quad (5)$$

przy czym wskaźnik W_j^s obliczany jest według (4).

2. Wyniki badań

Respondentami w badaniu gustów muzycznych byli użytkownicy portalu internetowego YouTube. Odpowiednią ankietę zawierającą między innymi pytania o ulubiony gatunek muzyczny, epokę muzyczną, kompozytora oraz pytania o dane społeczno-demograficzne dotyczące respondentów rozesłano drogą elektroniczną do ok. 450 osób. Zwrot ankiet był na poziomie 50%, a po ostatecznej selekcji nadesłanych odpowiedzi w analizie uwzględniono 200 poprawnie wypełnionych ankiet. Respondenci wybierali hipotetyczny wariant „produktu muzycznego”, którym może być np. płyta kompaktowa z nagraniami

muzyki różnych kompozytorów lub koncert w filharmonii. Przy ocenie tego rodzaju produktu pod uwagę wzięto dziesięć zmiennych (tabela 3), dlatego w badaniu każdy z ocenianych wariantów traktowano jako wielowymiarowy obiekt.

Tabela 3

Zmienne oceniające preferencje muzyczne

Nr zmiennej	Zmienna /atrybut/	Poziom zmiennej
1	Ulubiona epoka muzyczna	barok [A]
		klasycyzm [B]
		romantyzm [C]
2	Preferowany gatunek muzyki wokально-instrumentalnej	opera [A]
		oratorium [B]
		msza [C]
3	Preferowany gatunek muzyki orkiestrowej	symfonia [A]
		koncert [B]
4	Ulubiony kompozytor epoki baroku	Bach [A]
		Haendel [B]
		Vivaldi [C]
5	Ulubiony kompozytor epoki klasycyzmu	Mozart [A]
		Beethoven [B]
		Haydn [C]
6	Ulubiony kompozytor epoki romantyzmu	Brahms [A]
		Schumann [B]
		Chopin [C]
7	Preferowany kompozytor muzyki XX wieku	Strawiński [A]
		Bartok [B]
		Debussy [C]
8	Ulubiony twórca oper	Wagner [A]
		Verdi [B]
		Puccini [C]
9	Preferowany nurt muzyki XX wieku	neoklasycyzm [A]
		ekspresjonizm [B]
		impresjonizm [C]
10	Programowość muzyki	tak [A]
		nie [B]

Na podstawie dziesięciu wyróżnionych zmiennych w ankiecie oraz odpowiadających im poziomów utworzono zbiór hipotetycznych wariantów wyboru profilów. Spośród 26 244 profilów zaproponowano 21 wariantów, które poddano ocenie respondentów (tabela 4). Konieczność redukcji liczby profilów wynikała z faktu, iż ankietowane osoby nie będą w stanie ocenić 26 244 wariantów.

Tabela 4

Profile służące do oceny preferencji muzycznych

Profil	Czynniki określające preferencje muzyczne									
	1	2	3	4	5	6	7	8	9	10
1	A	B	A	B	A	B	B	B	A	B
2	A	A	B	C	B	A	C	C	B	B
3	A	C	A	A	B	C	C	B	C	A
4	A	A	A	B	C	B	A	A	A	A
5	A	B	B	A	B	A	C	A	A	B
6	A	C	A	A	A	C	B	A	B	B
7	A	A	B	A	C	B	C	C	C	A
8	B	C	A	C	B	A	B	A	A	B
9	B	A	B	B	A	C	A	A	B	A
10	B	A	A	C	B	A	C	B	C	A
11	B	C	A	A	C	B	B	B	A	B
12	B	B	A	C	C	A	A	B	C	A
13	B	B	A	B	B	C	B	B	A	A
14	B	C	B	B	A	B	C	C	B	A
15	C	A	B	A	B	A	B	C	A	A
16	C	B	A	C	C	B	C	A	A	A
17	C	C	B	A	B	C	A	C	B	A
18	C	B	B	A	A	A	A	A	A	B
19	C	A	B	C	A	B	C	A	B	A
20	C	A	A	A	A	A	B	B	B	B
21	C	B	B	A	B	A	C	C	A	B

Ankietowani melomani porządkowali wyróżnione profile zgodnie ze swoimi preferencjami: od najmniej preferowanego do najbardziej preferowanego przypisując im rangi od 1 do 21. Zmienna Y_k w modelu (1) jest więc mierzona na skali porządkowej. Na podstawie ocen profili dokonanych przez respondentów przeprowadzono estymację parametrów modelu (1) metodą najmniejszych kwadratów. Czynniki przedstawione w tabeli 2 zostały uwzględnione w modelu z pomocą zmiennych sztucznych zgodnie z procedurą kodowania quasi-eksperymentalnego. Oszacowanie parametrów modelu (1) dla każdego z respondentów umożliwiło obliczenie użyteczności częściowych na podstawie formuł zawartych w tabelach 1 i 2.

Wyniki użyteczności częściowych respondentów oraz średnie użyteczności częściowe w analizie conjoint przedstawiono w tabeli 5.

Tabela 5

Wyniki użyteczności cząstkowych w analizie conjoint

Atrybut	Poziom zmiennej	Nr respondenta					Średnia
		1	2		199	200	
Ulubiona epoka muzyczna	barok	-0,438	0,383		1,477	3,581	1,997
	klasycyzm	-2,733	1,408		-0,800	-0,765	-4,450
	romantyzm	3,171	-1,791		-0,677	-2,815	2,453
Preferowany gatunek muzyki wokально-instrumentalnej	opera	-1,523	1,969		1,808	-1,071	-2,597
	oratorium	-1,803	-2,552		-3,115	1,461	-1,120
	msza	3,327	0,582		1,307	-0,390	3,717
Preferowany gatunek muzyki orkiestrowej	symfonia	-1,874	1,530		2,293	0,251	-3,094
	koncert	1,874	-1,530		-2,293	-0,251	3,094
Ulubiony kompozytor epoki baroku	Bach	3,505	6,729		5,378	6,205	0,527
	Haendel	0,284	-6,841		-4,105	-7,148	-0,004
	Vivaldi	-3,789	0,112		-1,273	0,943	-0,522
Ulubiony kompozytor epoki klasycyzmu	Mozart	1,499	8,995		8,262	6,298	3,861
	Beethoven	-2,625	-6,554		-4,632	0,062	-3,069
	Haydn	1,126	-2,441		-3,630	-6,360	-0,793
Ulubiony kompozytor epoki romantyzmu	Brahms	1,322	-1,608		-0,886	-1,053	1,655
	Schumann	-4,466	-5,239		-4,274	-0,104	-5,713
	Chopin	3,144	6,847		5,159	1,157	4,058
Preferowany kompozytor muzyki XX wieku	Strawiński	-0,196	2,717		3,822	3,265	0,756
	Bartok	-0,119	-6,370		-5,845	-1,891	0,027
	Debussy	0,315	3,654		2,023	-1,374	-0,783
Ulubiony twórca oper	Wagner	-2,563	-7,172		-5,888	-6,946	-5,356
	Verdi	2,906	2,770		1,346	3,557	5,764
	Puccini	-0,343	4,402		4,542	3,388	-0,408
Preferowany nurt muzyki XX wieku	neoklasycyzm	2,550	11,081		9,466	6,043	4,814
	ekspresjonizm	-3,253	-4,992		-5,469	-0,024	-2,889
	impresjonizm	0,703	-6,089		-3,997	-6,019	-1,925
Programowość muzyki	tak	0,302	0,495		0,820	2,568	0,905
	nie	-0,302	-0,495		-0,820	-2,568	-0,905

Na podstawie średnich wartości użyteczności cząstkowych zawartych w tabeli 5 można stwierdzić, iż respondenci na ogół preferują muzykę epoki romantycznej, a w dalszej kolejności muzykę barokową, natomiast mniejszym uznaniem cieszy się muzyka epoki klasycyzmu. Spośród wyróżnionych gatunków muzycznych najbardziej preferowane są msza i koncert, natomiast uznania nie znalazły opera, oratorium oraz symfonia. Respondenci najchętniej gustują w muzyce Bacha, Mozarta, Brahmsa, Chopina, Strawińskiego oraz Bartoka. Z kolej mniejszą popularnością cieszy się muzyka komponowana

przez Haendla, Vivaldiego, Beethovena, Haydna, Schumana i Debussyego. Spośród wyróżnionych twórców oper najchętniej wybierany był Verdi, natomiast uznania nie znalazła operowa muzyka Wagnera i Pucciniego. Ankietowani melomani preferują muzykę neoklasyczną, natomiast mniejszym zainteresowaniem cieszy się u nich muzyka ekspresjonistyczna i impresjonistyczna. Typowy respondent częściej wybiera również muzykę programową niż absolutną (pozbawioną treści pozamuzycznych). W tabeli 6 przedstawiono wyniki relatywnej ważności atrybutów dla poszczególnych respondentów oraz średnią relatywną ważność zmiennych obliczone według wzorów (4) i (5).

Tabela 6

Relatywna ważność atrybutów (w %) w analizie conjoint

Atrybut	Nr respondenta					Średnia
	1	2		199	200	
Ulubiona epoka muzyczna	12,78	3,49		2,84	9,07	11,63
Preferowany gatunek muzyki wokально-instrumentalnej	2,84	4,93		6,13	3,59	10,64
Preferowany gatunek muzyki orkiestrowej	6,13	3,33		5,71	0,71	10,43
Ulubiony kompozytor epoki baroku	5,71	14,79		11,81	18,94	1,77
Ulubiony kompozytor epoki klasycyzmu	11,81	16,95		16,06	17,95	11,68
Ulubiony kompozytor epoki romantyzmu	16,06	13,17		11,75	3,13	16,47
Preferowany kompozytor muzyki XX wieku	11,75	10,93		12,04	7,31	2,59
Ulubiony twórca oper	12,04	12,62		12,99	14,90	18,74
Preferowany nurt muzyki współczesnej	12,99	18,72		18,61	17,11	12,98
Programowość muzyki	18,61	1,08		2,04	7,28	3,05

Z rezultatów przedstawionych w ostatniej kolumnie tabeli 6 wynika, iż największe znaczenie wśród atrybutów reprezentujących gust, estetykę muzyczną i kulturę muzyczną respondentów ma ulubiony twórca opery, a w dalszej kolejności ulubiony kompozytor epoki romantyzmu. Stosunkowo niewielkie znaczenie w rozpoznaniu preferencji muzycznych melomanów okazały się mieć ulubiony kompozytor epoki baroku, preferowany kompozytor muzyki XX wieku oraz programowość muzyki.

Wartości całkowitej atrakcyjności ocenianych profili poszczególnych respondentów i średnie wartości całkowitej atrakcyjności profili obliczone na podstawie wzorów (2) i (3) przedstawiono w tabeli 7.

Tabela 7

Całkowita atrakcyjność ocenianych profili

Numer porofilu	Nr respondenta					Średnia
	1	2		199	200	
1	9,415	11,584		13,608	17,453	11,166
2	2,415	3,666		3,400	9,608	9,066
3	20,443	18,669		20,097	17,571	13,895
4	4,380	4,804		10,710	2,052	7,185
5	13,012	10,260		9,633	13,134	11,994
6	14,103	14,360		14,777	13,647	15,090
7	12,233	10,656		10,775	8,535	10,704
8	4,372	0,837		1,846	1,660	9,001
9	8,012	10,219		10,237	5,056	10,568
10	4,182	6,010		5,625	5,073	7,112
11	15,099	17,877		13,345	11,951	12,799
12	7,142	4,664		3,503	5,821	8,839
13	11,210	10,136		9,510	13,268	13,067
14	7,984	9,257		8,935	10,170	10,615
15	19,292	15,147		16,602	19,157	15,754
16	4,146	5,999		4,667	1,640	8,074
17	20,084	15,228		16,879	21,138	13,914
18	20,234	22,697		22,171	17,612	12,930
19	2,743	2,824		1,960	5,198	8,208
20	16,510	15,059		14,312	14,863	11,023
21	18,842	19,659		17,909	17,071	11,332

Na podstawie średnich wartości całkowitej atrakcyjności profili można stwierdzić, iż najwyżej były oceniane profile 15 i 6, a najmniejsze znaczenie dla respondentów miał profile 4 i 10.

Cząstkowe użyteczności respondentów (tabela 6) odzwierciedlają ich reakcje na poszczególne oceniane profile [10], co pozwala wykorzystać te użyteczności do przeprowadzenia segmentacji melomanów według podobnych gustów i preferencji muzycznych. W tym celu wykorzystano metodę *k*-średnich, przy czym zdecydowano się na wyodrębnienie czterech jednorodnych grup melomanów. Analiza conjoint została następnie przeprowadzona osobno w każdym z wyróżnionych segmentów melomanów. Aby określić zdolność poziomów atrybutów do przeprowadzenia segmentacji respondentów, wykorzystano jednoczynnikową analizę wariancji ANOVA. Przeciętne wartości użyteczności cząstkowych respondentów w poszczególnych segmentach wraz z wynikami analizy wariancji przedstawiono w tabeli 8.

Tabela 8

Wyniki segmentacji grupy respondentów

Atrybut	Poziom zmiennej	Numer segmentu				F	Wartość <i>p</i>
		I	II	III	IV		
		średnia użyteczność cząstkowa					
Ulubiona epoka muzyczna	barok	1,084	1,029	3,113	2,847	1,221	0,3031
	klasycyzm	-1,832	2,472	5,735	-3,722	6,153	0,0005
	romantyzm	0,749	-3,501	-8,848	0,875	6,661	0,0003
Preferowany gatunek muzyki wokaln- -instrumentalnej	opera	-1,208	-0,436	1,070	-1,131	0,929	0,4279
	oratorium	-0,258	0,343	2,676	-4,152	7,174	0,0001
	msza	1,466	0,093	-3,746	5,283	7,552	0,0001
Preferowany gatunek muzyki orkiestrowej	symfonia	-0,964	2,026	4,815	-8,304	9,662	0,0000
	koncert	0,964	-2,026	-4,815	8,304	9,662	0,0000
Ulubiony kompozytor epoki baroku	Bach	4,049	-2,628	5,490	-2,261	4,634	0,0037
	Haendel	-2,468	-1,741	-8,947	0,366	3,924	0,0095
	Vivaldi	-1,580	4,369	3,457	1,895	4,059	0,0079
Ulubiony kompozytor epoki klasycyzmu	Mozart	2,966	-0,119	6,974	-1,028	2,075	0,1049
	Beethoven	-1,740	0,846	0,391	-8,100	3,627	0,0140
	Haydn	-1,225	-0,727	-7,365	9,128	11,176	0,0000
Ulubiony kompozytor epoki romantyzmu	Brahms	0,202	-1,195	-0,145	1,241	1,421	0,2380
	Schumann	-2,905	2,073	2,503	-10,018	8,120	0,0000
	Chopin	2,703	-0,877	-2,358	8,777	6,728	0,0002
Preferowany kompo- zytor muzyki XX wieku	Strawiński	0,565	1,597	3,259	-7,818	11,176	0,0000
	Bartok	-0,549	1,164	-2,280	6,062	3,590	0,0147
	Debussy	-0,016	-2,762	-0,979	1,756	1,647	0,1800
Ulubiony twórca oper	Wagner	-3,840	2,300	-1,234	-2,123	2,441	0,0656
	Verdi	3,478	-5,227	-3,860	5,294	8,135	0,0000
	Puccini	0,362	2,927	5,094	-3,171	3,368	0,0197
Preferowany nurt muzyki XX wieku	neoklasycyzm	4,339	-1,731	5,031	4,330	0,779	0,5071
	ekspresjonizm	-1,185	2,324	1,767	-3,390	2,854	0,0384
	impresjonizm	-3,153	-0,593	-6,798	-0,940	1,696	0,1692
Programowość muzyki	tak	1,157	1,526	4,707	0,840	3,642	0,0137
	nie	-1,157	-1,526	-4,707	-0,840	3,642	0,0137

Z przedstawionych rezultatów w tabeli 8 wynika, iż większość poziomów atrybutów (20 na 28 przypadków) statystycznie istotnie różnicowała powstałe segmenty respondentów (prawdopodobieństwa testowe im odpowiadające były mniejsze od 0,05). Przeciętne użyteczności według wyróżnionych segmentów

pozwalają na sprawdzenie w ramach każdego atrybutu, który z poziomów tego atrybutu był najbardziej preferowany, a który był najmniej preferowany. Na przykład muzyka epoki klasycyzmu cieszy się największym uznaniem w segmentach trzecim i drugim, natomiast nie znalazła uznania w pozostałych grupach melomanów.

Średnią ważność zmiennych według wyróżnionych segmentów przedstawiono w tabeli 9.

Tabela 9

Średnia ważność zmiennych według segmentów respondentów

Atrybut	Numer segmentu			
	I	II	III	IV
	średnia wartość zmiennych (w %)			
Ulubiona epoka muzyczna	8,931	10,869	15,297	7,778
Preferowany gatunek muzyki wokalnie-instrumentalnej	7,626	1,872	6,674	8,396
Preferowany gatunek muzyki orkiestrowej	6,330	7,350	9,468	13,364
Ulubiony kompozytor epoki baroku	15,454	14,360	14,488	6,818
Ulubiony kompozytor epoki klasycyzmu	11,731	8,538	12,367	15,367
Ulubiony kompozytor epoki romantyzmu	10,114	11,355	5,061	14,662
Preferowany kompozytor muzyki XX wieku	7,701	9,758	6,098	11,925
Ulubiony twórca oper	13,457	18,850	11,622	8,215
Preferowany nurt muzyki współczesnej	14,441	11,337	10,581	10,680
Programowość muzyki	4,216	5,711	8,343	2,794

W pierwszej grupie melomanów największe znaczenie w kształtowaniu preferencji muzycznych mają twórcy muzyki okresu baroku, a w dalszej kolejności nurty muzyki współczesnej. W drugim segmencie melomanów szczególną wagę respondenci przywiązują do twórców oper i kompozytorów muzyki barokowej. Z kolei w trzeciej grupie ankietowanych melomanów największe znaczenie miał wybór epoki muzycznej oraz wybór kompozytora muzyki barokowej. Natomiast w czwartym segmencie respondentów największe znaczenie mają twórcy muzyki okresu klasycyzmu i romantyzmu. Tabela 10 zawiera średnie wartości całkowitej atrakcyjności profilów według wyróżnionych segmentów melomanów.

Tabela 10

Całkowita atrakcyjność ocenianych profili według segmentów respondentów

Numer profilu	Numer segmentu			
	I	II	III	IV
	średnia wartość zmiennych			
1	13,283	7,932	11,552	7,535
2	5,444	15,193	10,478	12,388
3	17,782	5,075	7,007	18,171
4	4,253	17,556	13,186	1,093
5	13,343	4,292	11,055	13,980
6	14,290	15,427	14,067	18,001
7	8,827	10,025	8,254	18,331
8	4,061	20,459	13,550	8,700
9	7,541	16,661	12,392	12,054
10	4,061	12,668	14,625	1,807
11	14,417	7,629	7,849	17,930
12	6,108	16,232	12,714	6,441
13	13,584	10,442	12,143	14,370
14	8,228	16,410	14,527	8,198
15	18,044	6,588	11,929	19,967
16	5,173	15,558	10,997	7,205
17	18,810	11,924	7,175	12,319
18	18,296	3,156	9,917	9,508
19	4,819	15,391	10,836	8,958
20	16,098	2,523	6,509	9,496
21	17,211	0,389	5,422	10,961

Największą atrakcyjnością w pierwszej grupie respondentów cieszyły się profile 17 i 18. Z kolei w drugiej grupie respondentów największe znaczenie miały profile 8 i 4. Trzeci segment ankietowanych melomanów za najbardziej atrakcyjne uważa profile 10 i 14. Natomiast w ostatnim segmencie respondentów za najbardziej atrakcyjne uchodzą profile 15 i 7.

Aby dokładniej scharakteryzować wyróżnione segmenty respondentów, zbadano je pod względem wybranych cech społeczno-demograficznych, a mianowicie: płci, wieku, wykształcenia, miejsca zamieszkania, sposobu słuchania muzyki, tradycji muzycznych w rodzinie, uczestnictwa w wydarzeniach muzycznych (tabela 11).

Tabela 11

Wybrane cechy społeczno-demograficzne według
segmentów respondentów

Charakterystyka segmentu		Numer segmentu			
		I	II	III	IV
Średni wiek (w latach)		30,5	25,8	33,5	27
Średnia częstotliwość uczestniczenia w muzycznych wydarzeniach artystycznych w ciągu miesiąca		2,1	1,8	2,8	3,1
Średnia miesięczna kwota, jaką respondenci są skłonni przeznaczyć na cele kulturalne związane z muzyką (w euro)		52,4	72,1	46,2	101,2
udział w %					
Płeć	kobiety	56,8	42,1	51,7	46,2
	mężczyźni	43,2	57,9	48,3	53,8
Miejsce zamieszkania	miasto	91,7	86,4	92,6	86,1
	wieś	8,3	13,6	7,4	13,9
Wykształcenie	wyższe	72,4	60,4	71,6	74,8
	średnie	18,7	29,7	22,5	20
	podstawowe	8,9	9,9	5,9	5,2
Wykształcenie muzyczne	tak	16,7	18,2	22,7	9,6
	nie	83,3	81,8	77,3	90,4
Muzyczne tradycje rodzinne	tak	21,8	33,9	32,1	18,5
	nie	78,2	66,1	67,9	81,5
Zawód związany z muzyką	tak	14,1	16,1	17,2	8,2
	nie	85,9	83,9	82,8	91,8
Preferowana forma słuchania muzyki	sala koncertowa, operowa	28,6	31,4	39,8	43,6
	radio lub telewizja	11,9	4	5,2	17,3
	płyty kompaktowe	42	51,2	27,4	29,8
	Internet	17,5	13,4	27,6	9,3

Na podstawie wyników w tabeli 11 można stwierdzić, że średni wiek respondentów w żadnym z wyróżnionych segmentów nie przekraczał 35 lat. Melomani tworzący segment czwarty najczęściej uczestniczą w muzycznych wydarzeniach artystycznych i jednocześnie są skłonni przeznaczyć największe środki pieniężne na cele kulturalne związane z muzyką. Wśród ankietowanych osób udział kobiet jest nieco większy w segmentach pierwszym i trzecim, a mężczyźni przeważają w grupach drugiej i czwartej. Zdecydowana większość respondentów w każdym z wyróżnionych segmentów mieszka w mieście i posiada wykształcenie wyższe. Ponadto dominujący udział we wszystkich grupach mają osoby bez wykształcenia muzycznego, bez rodzinnych tradycji muzycznych i wykonujące zawód niezwiązany z muzyką. W segmentach pierw-

szym i drugim dominują osoby, które preferują słuchanie muzyki z płyt kompaktowych, natomiast w segmentach trzecim i czwartym największy udział mają osoby, dla których sala koncertowa czy opera są właściwym miejscem obcowania z muzyką.

Podsumowanie

Analiza conjoint została stworzona z myślą o prowadzeniu badań marketingowych i w tym obszarze znajduje najwięcej zastosowań. W niniejszym opracowaniu wykorzystano ją do oceny preferencji muzycznych. Podjęto tu próbę kwantyfikacji trudno mierzalnych kategorii, na które składają się między innymi piękno i harmonia brzmienia muzyki, artyzm języka dźwiękowego, głębia wyrazu dzieła muzycznego czy ponadczasowość treści przekazywanych poprzez muzykę. Ich ocena zależy między innymi od cech osobowościowych respondenta, jego wrażliwości muzycznej, wcześniejszego przygotowania edukacyjnego z zakresu muzyki i innych. Dlatego też melomani stanowią grupę w znacznym stopniu niejednorodną pod względem upodobań muzycznych.

Połączenie wyników analizy conjoint z metodą grupowania *k*-średnich pozwoliło na wyodrębnienie jednorodnych pod względem preferencji segmentów respondentów i szczegółową prezentację ich gustów i preferencji muzycznych. Wydaje się, że otrzymane wyniki mogą być interesujące zarówno dla psychologów, socjologów, jak i osób związanych ze środowiskiem muzycznym. Do pewnego stopnia mogą być również pomocne w dalszych badaniach służących dostosowaniu oferty uczestników rynku muzycznego do oczekiwań melomanów.

Literatura

1. Bąk A. (2004). Mikroekonometryczne metody badania preferencji. Zeszyty Naukowe. Wydawnictwo Uniwersytetu Szczecińskiego, Szczecin, 394.
2. Białkowski A., Grusiewicz M. (2009). Zainteresowania muzyczne i główne formy ich realizowania przez uczniów. Wychowanie muzyczne w Szkole, 3, UMCS, Lublin.
3. Ginocchio J. (2008). Music Style Preference: A Ranking of Musical Styles and Comparison by Age; Gender; Music Training; and Rural, Suburban, or Urban Upbringing. VDM Verlag.
4. Green P.E., Wind Y. (1975). New Way to Measure Consumers' Judgments. Harvard Business Review, 53, s. 107-117.
5. Hair J.F., Anderson R.E., Tatham R.L., Black W.C. (1995). Multivariate Data Analysis with Readings. Englewood Cliffs, Prentice-Hall.

6. Hallam S., Coss I., Thaut M. (2009). *Music Handbook of Music Psychology*. Oxford University Press.
7. Hargreaves D.J., North A.C. (1997). *The Social Psychology of Music*. Oxford University Press, Oxford.
8. Kamińska B. (2002). Upodobania muzyczne-problemy i wyniki badań. W: *Blżej muzyki, blżej człowieka*. Red. B. Białkowski, B. Smoleńska-Zielińska. UMCS, Lublin.
9. Lonsdale A.J., North A.C. (2011). Why do we Listen to Music? A Uses and Gratifications Analysis. *British Journal of Psychology*, 102, s. 108-134.
10. Walesiak M., Bąk A. (2000). *Conjoint analysis w badaniach marketingowych*. Wydawnictwo Akademii Ekonomicznej, Wrocław.
11. Walesiak M. (1996). *Metody analizy danych marketingowych*. PWN, Warszawa.
12. Waldorff J. (1982). *Muzyka łagodzi obyczaje*. Polskie Wydawnictwo Muzyczne, Kraków.

AN APPLICATION OF CONJOINT ANALYSIS TO STUDY MUSIC PREFERENCES

Summary

The aim of the paper is to study the music preferences of the music lovers in area of the type of classical music, favorite composers and eras of music. The basic research method used in the paper is the conjoint analysis. Its application allows to estimate the respondents' utilities, which enabled calculating the relative importance of each variable representing musical taste, aesthetics of music and musical culture of the test group of the music lovers. In addition, estimates of the partial utility values, led to segmentation of the music lovers by similar preferences. The implementation of conjoint analysis in each of the segments allowed to present detailed description of the music preferences of the respondents. The results of research may be helpful in adapting the music publishing deals, publishing labels, repertoire of concert halls and music programs on radio and television to the needs of the music lovers.

Andrzej M.J. Skulimowski

AGH Akademia Górniczo-Hutnicza w Krakowie

METODY PROGNOZOWANIA DYNAMICZNYCH RANKINGÓW PRIORYTETÓW TECHNOLOGICZNYCH*

Wprowadzenie

W niniejszym opracowaniu przedstawimy podstawy teoretyczne oraz wynikające z nich praktyczne algorytmy rozwiązywania istotnej klasy wielokryterialnych problemów decyzyjnych, występujących przy podejmowaniu decyzji inwestycyjnych w oparciu o rankingi celów i priorytetów. Metodyka określania priorytetów społecznych, gospodarczych, ekologicznych i technologicznych w kontekście prognoz i scenariuszy uzyskanych jako rezultaty badań foresightowych wymaga zdefiniowania (najczęściej w kontekście celów średnio-terminowych) wskaźników ilościowych oraz funkcji estymujących korzyści socjalne, gospodarcze i ekologiczne. Następnie konieczne jest uwzględnienie aktualnych warunków ekonomicznych wdrażania strategii związanych z rozwojem technologicznym i wynikających z nich ograniczeń dla parametrów finansowych i ekonomicznych w rozpatrywanym horyzoncie czasowym prognozy. Takie podejście prowadzi do sformułowania wielokryterialnego zagadnienia decyzyjnego jako problemu optymalizacji wektorowej ze zbiorem kryteriów opisujących socjalne, ekonomiczne, ekologiczne, jak również technologiczne preferencje interesariuszy danego przedsięwzięcia [18]. Równolegle formułowany jest pomocniczy problem wielokryterialnego uszeregowania obiektów [19] ze względu na powiązany z powyższymi preferencjami zbiór kryteriów. Zakładamy przy tym, że wyniki optymalizacji rozwiązania problemu nadrzędnego są zależne od wyboru portfela obiektów, takich jak projekty tech-

* Artykuł został opracowany na podstawie badań wykonanych w ramach projektów pt. „Scenariusze i trendy rozwojowe wybranych technologii społeczeństwa informacyjnego do roku 2025”, współfinansowanego ze środków Unii Europejskiej – Poddziałania 1.1.1 Programu Operacyjnego Innowacyjna Gospodarka, lata 2007-2013, nr projektu WND-POIG.01.01.01-00-021/09 oraz „Odpady nieorganiczne przemysłu chemicznego – foresight technologiczny” finansowanego w ramach Poddziałania 1.1.1 POIG, nr projektu WND-POIG.01.01.01-00-009/09.

nologiczne czy infrastrukturalne, na podstawie tego uszeregowania, [7; 10]. W szczególności, wybór priorytetowych technologii proekologicznych, których wdrażanie w przedsiębiorstwach wspierane jest ze środków Programów Operacyjnych, ma bezpośredni wpływ na wartości kryteriów ekologicznych, ekonomicznych i społecznych charakteryzujących region, w którym realizowane są te przedsięwzięcia.

Nawiązując do struktury problemu i terminologii stosowanej dalej, dla analizowanej tu klasy zastosowań decydentami są z reguły centralne urzędy branżowe lub władze regionu, lista priorytetów zawarta jest w narodowej lub odpowiednio – regionalnej strategii rozwoju, zgodnej z Narodową Strategią Spójności, a rankingi priorytetów, technologii kluczowych, projektów technologicznych i infrastrukturalnych tworzone są przez ekspertów i zatwierdzane przez decydentów. Natomiast interesariuszem jest ogół społeczeństwa rozważanego terytorium reprezentowany przez instytucje i organizacje nadzorujące procesy alokacji wsparcia. Szczególnie ważna jest umiejętność zastosowania metod wielokryterialnych w procesach decyzyjnych związanych z prognozowaniem efektywnego rozdziału środków finansowych na poszczególne obszary polityki naukowo-badawczej lub rozwojowej.

Przypomnijmy, że problem optymalizacji wielokryterialnej może być sformułowany następująco

$$(F: U \rightarrow E) \rightarrow \min(\theta) \quad (1)$$

gdzie U i E to odpowiednio przestrzeń decyzji i przestrzeń kryteriów tj. zbiór, w którym wartości przyjmuje funkcja kryterialna F , $F = (F_1, \dots, F_n)$ jest wektorową funkcją celu, θ jest domkniętym i wypukłym stożkiem, wprowadzającym częściowy porządek w E [5; 13]. Porządek ten będzie oznaczany jako $\succsim_\theta$, z definicji

$$x \succsim_\theta y \Leftrightarrow y - x \in \theta$$

W najczęstszym przypadku, analizowanym również i w tym opracowaniu, $E = \mathbb{R}^n$, a θ jest dodatnim stożkiem $\mathbb{R}_+^n$, lub θ jest izomorficzny z $\mathbb{R}_+^n$ i $\eta(\mathbb{R}_+^n) \subset \theta$ dla pewnej izometrii η . Ten ostatni warunek gwarantuje istnienie takich kombinacji liniowych $G_1, \dots, G_n$ kryteriów $(F_1, \dots, F_n)$, że problem (1) jest równoważny problemowi $(G: U \rightarrow E) \rightarrow \min$, gdzie minimum określone jest przez naturalny stożek $\mathbb{R}_+^n$, któremu odpowiada naturalny porządek częściowy „ $\leq$ ” wyznaczony przez porównanie kolejnych współrzędnych wektorów x i y , tj.

$$x \leq y \Leftrightarrow x_i \leq y_i \text{ dla } i=1, \dots, n$$

Zbiór punktów niezdominowanych w problemie (1) będzie oznaczany jako $P(U, F, \theta)$, z definicji $x \in P(U, F, \theta)$ wtedy i tylko wtedy, gdy

$$\forall y \in U [y \leq x \Rightarrow x = y]$$

W pewnych sytuacjach problem znalezienia rankingu elementów zbioru X_0 może zostać sformułowany jako sekwencja wielokryterialnych problemów decyzyjnych typu (1)

$$(F: X_k \rightarrow \mathbb{R}^n) \rightarrow \min(\theta_k) \quad (2)$$

$$(\varphi_k: P(X_k, F_k, \theta_k) \rightarrow \mathbb{R}) \rightarrow \min \quad (3)$$

$$X_k := X_{k-1} \setminus \operatorname{argmin}\{(\varphi_{k-1} \circ F_{k-1})(X_{k-1})\} \quad (4)$$

dla $k = 1, \dots, K-1$, gdzie (3) jest problemem optymalizacji wielokryterialnej typu (1), $X_0 := X$ jest zbiorem alternatyw o liczności większej od K , $F_0 := F = (f_1, \dots, f_n), \dots, F_k = (f_{1,k}, \dots, f_{n,k})$ są wektorami funkcji kryterialnych dla k -tego problemu, $k = 1, \dots, K$, a θ_k jest stożkiem definiującym porządek częściowy w $\mathbb{R}^n$. Każdy z problemów (2)-(3) opisuje preferencje interesariuszy, jak również stopień osiągnięcia celów osób podejmujących decyzje.

Problem szeregowania spełniający warunek (4) będzie nazywany bezkontekstowym. W problemach portfelowych tego typu wybór jednego z obiektów do portfela jest niezależny od wyboru dowolnego z pozostałych obiektów [17]. Dla zapewnienia obiektywności tworzenia rankingu założymy dodatkowo, że

$$\forall k \in \{1, \dots, K\}, F_k := F, \varphi_k := \varphi \text{ na } X_0 \text{ i } \theta_k = \theta$$

Gdy U jest zbiorem ocenianych technologii, wtedy rozwiązaniem problemu (1) (a właściwie ciągu takich problemów) może być ranking priorytetów technologicznych, które mogą wchodzić w skład różnych strategii. Aby ranking priorytetów uczynić możliwym do wdrożenia praktycznego, należy precyzyjnie zdefiniować zależność pomiędzy metodyką ranking, a celami praktycznymi. Tak zdefiniowane priorytety są zwykle stosowane w celu przydziału zasobów dla różnych przedsięwzięć badawczych lub wdrożeniowych. Innym zastosowaniem analizy priorytetów jest wykorzystanie przez decydentów rankingu do wskazania hierarchii czasowej tak, aby cele scharakteryzowane przez priorytety o wyższym rankingu były realizowane wcześniej niż te o niższym priorytecie.

Wadą dotychczas stosowanych podejść jest statyczny model priorytetów, w którym początkowe rangi obiektów, np. technologii kluczowych zostają określone dla całego okresu wdrażania, przy czym brakuje zobiektywizowanego mechanizmu ich modyfikacji w zależności od zmieniającej się sytuacji zewnętrznej (gospodarczej, ekologicznej, czy socjalnej) lub w przypadku wcześniejszego od zakładanego osiągnięcia określonych celów strategicznych. Problemy wynikające z uwzględnienia dynamiki przy określaniu priorytetów rozważane były także w [1; 20], jednakże bez wzięcia pod uwagę możliwości ewolucji raz ustalonych rankingów w wyniku zmieniających się preferencji i okoliczności. W praktyce nowe priorytety technologiczne i odpowiadające im strategie

wdrożeniowe redefiniowane są najczęściej po zmianie decydentów. Wtedy też tworzone są nowe prognozy i scenariusze, często bez powiązania z wcześniejszymi badaniami. Ponadto istniejące modele przyszłości stosowane w procesach wspomagania decyzji technologicznych nie uwzględniają na ogół w wystarczający sposób powiązań pomiędzy celami dotąd osiągniętymi, a tymi, które mają być osiągnięte w przyszłości, wskutek czego rozbieżności pomiędzy prognoząmi a rzeczywistością często zaskakują decydentów.

Innowacyjność metody dynamicznych rankingów priorytetów przedstawionej w dalszej części niniejszego opracowania polega także na tym, że wzajemne wpływy prognozy scenariuszowej oraz czynników zewnętrznych wpływających na przyszłe scenariusze są analizowane w jednym wspólnym modelu. Wzajemne zależności między zdarzeniami mogą być modelowane jako sterowany system zdarzeń dyskretnych, dzięki czemu sieć przyszłych zdarzeń i przyczynowych relacji między nimi można przedstawić jako graf o zmiennej strukturze zależnej od czasu. Wierzchołki grafu interpretowane są jako stany rozważanych obiektów, parametryzowane przez wartości optymalizowanych kryteriów.

1. Dynamiczne rankingi a foresight

Metodyka oparta na wynikach badań foresightowych przy określaniu obiektywnego algorytmu przydziału zasobów jest szeroko akceptowanym podejściem wspierającym podejmowanie strategicznych decyzji na różnych szczeblach [15; 18]. Jednym z możliwych celów prognozowania scenariuszowego jest utworzenie rankingu celów strategicznych i zarekomendowanie ich osobom podejmującym decyzje. Horyzont czasowy foresightu jest relatywnie długi i zawiera się zazwyczaj w przedziale 15-50 lat [18]. Zakładamy, iż priorytety technologiczne są rezultatem prognoz ekonomiczno-finansowych, społecznych i technologicznych i że przedstawione są one w formie rekomendacji i rankingów. Należy zaznaczyć, iż odmiennie jak to jest w przypadku indywidualnych rankingów projektów priorytet technologiczny, nie jest jednoznacznie zdefiniowanym terminem, zatem pojęcie to musi zostać powiązane z planem konkretnych działań wdrożeniowych, ekonomicznych lub ekologicznych.

Powyższe ustalenia prowadzą do sformułowania hierarchicznego wielokryterialnego problemu decyzyjnego, który uwzględnia wyniki prognoz i badań foresightowych za pomocą optymalizacji kryteriów opisujących społeczne, ekonomiczne, ekologiczne i technologiczne preferencje zarówno interesariuszy, jak i decydentów. Pod uwagę brany jest również stopień osiągnięcia celów strategicznych zatwierdzonych przez władze legislacyjne i będących elementem polityki naukowo badawczej, rozwojowej lub regionalnej.

Hierarchiczna procedura przydziału funduszy na wdrożenie priorytetów technologicznych może zostać scharakteryzowana przez następujące reguły:

1. Na najwyższym poziomie decyzyjnym dynamiczne rankingi są stosowane do szeregowania obszarów polityki.

2. Na niższych poziomach szeregowane są kolejno cele pośrednie i szczegółowe, a następnie technologie i projekty technologiczne, na które przydzielane są środki.

3. Uporządkowanie na poziomie niższym musi uwzględniać rankingi z wyższego poziomu, dodatkowe charakterystyki projektów (każdy projekt może być powiązany z kilkoma obszarami polityki) i wielkość funduszy przeznaczonych do przydziału na projekty (zbyt duże projekty mogą nie otrzymać finansowania pomimo wyższej pozycji w rankingu od projektów niżej ocenianych, lecz mieszczących się w limicie środków do podziału).

4. Każdy przydział środków traktowany jest jako funkcja czasu (np. kwartalny plan finansowania w okresie realizacji projektu) i zewnętrznych zmiennych logicznych, które reprezentują zmiany w rankingach wyższego poziomu i stanów zewnętrznego otoczenia politycznego, ekonomicznego, ekologicznego, społecznego i technologicznego (PEEST). Dodatkowo alokacja środków może być uzależniona od stopnia osiągnięcia celów strategicznych określonych na wyższych szczeblach podejmowania decyzji.

Charakterystyczną cechą przedstawionego podejścia, które należy do metodyki foresightu, jest adaptacyjne stosowanie scenariuszy do ustalenia przyszłych rankingów. Mianowicie, opierając się na estymacji prawdopodobieństw zdarzeń wpływających na spełnienie się scenariuszy w chwili t_0 , możemy wyliczyć prawdopodobieństwo poszczególnych scenariuszy, a poprzez zależność rang od wskaźników je opisujących, traktowanych jako kryteria, możliwa jest budowa wariantowych prognoz także i przyszłych rankingów. Pozwala to uwzględnić w planowaniu przydziału funduszy więcej wiedzy ex-ante niż w zazwyczaj stosowanych, statycznych podejściach, ograniczających się do analizy potrzeb w chwili początkowej.

Poza usprawnieniem finansowania projektów rozdzielanego w drodze konkursów, dynamiczne rankingi mogą być pomocne przy określaniu optymalnej struktury organizacyjnej i charakterystyk zasobów ludzkich instytucji wdrażających takie projekty. Prognoza kształtowania się przyszłych priorytetów może być też wykorzystana do planowania potrzeb budżetowych w przyszłości i działań, które należy podjąć w odpowiedzi na przewidywane zdarzenia zewnętrzne niezależne od decydenta. Rankingi obiektów wyższego poziomu mogą służyć do konstrukcji rankingów na niższych (operacyjnych) poziomach decyzyjnych, np. cele unijnej polityki ekologicznej czy preferowane kierunki ekspansji inwestycyjnej wpływają na rankingi technologii w przemyśle che-

micznym, które z kolei są uwzględniane w regionalnych programach operacyjnych przy rozdziale środków na projekty technologiczne przedsiębiorstw, obszary akcji prewencyjnych, strategii inwestycji komunalnych itp.

Przykład zastosowania opisanego tu podejścia do konstrukcji dynamicznych rankingów polityki regionalnej jest przedstawiony niżej.

2. Opis metody dynamicznych rankingów (DPR)

Zależności pomiędzy zdarzeniami zachodzącymi w rozpatrywanym systemie, sterowaniami wewnętrznymi, wpływem decyzji zewnętrznych i zdarzeń losowych mogą być modelowane jako sterowany automat o znanej funkcji przejścia, nazywany systemem zdarzeń dyskretnych [11]. Pozwalają tym samym na utworzenie modelu sieci przyszłych zdarzeń i relacji między nimi jako grafu o strukturze zależnej od czasu, gdzie wierzchołki interpretowane są jako stany obiektów, a krawędzie jako sterowania (czyli akcje podejmowane przez decydentów), z parametrami odpowiadającymi wartościom optymalizowanych kryteriów.

Przypomnijmy [11], że system zdarzeń dyskretnych może być przedstawiony jako układ

$$P = (Q, V, \delta, Q_0, Q_m) \quad (5)$$

gdzie: Q – skończony zbiór stanów systemu, V – zbiór dopuszczalnych operacji na stanach systemu (sterowań), $\delta: V \times Q \rightarrow Q$ – funkcja przejścia definiująca wyniki operacji na stanach, Q_0 – zbiór stanów początkowych analizowanych procesów, Q_m – zbiór (potencjalnych) stanów końcowych. Para stanów $e:=(q_1, q_2)$, takich, że $q_2=\delta(v, q_1)$ będzie nazywana zdarzeniem, co oznacza, że do stanu q_2 przechodzimy ze stanu q_1 na skutek wystąpienia zdarzenia (operacji) v ze zbioru V . Zbiór wszystkich dopuszczalnych zdarzeń w systemie (5) oznaczmy przez E . Zakładamy, że operacje ze zbioru V są albo konsekwencjami decyzji podejmowanych przez decydenta odpowiadającego za funkcjonowanie systemu (5), działaniami zewnętrznymi decydentów nad Q , albo że mogą pojawiać się spontanicznie, jako skutki losowych procesów. W pierwszym przypadku, operacje $v \in V$ będą nazywane sterowaniami. Zakładamy ponadto, że istnieje zbiór $X(Q)$, zawierający ilościowe bądź porządkowe charakterystyki zbioru stanów Q , który może być deterministyczny, stochastyczny, rozmyty itp. Podobnie, do zdarzeń może być przypisany zbiór kryteriów $G = (G_1, \dots, G_k): E \rightarrow \mathbb{R}^k$. Jedna ze współrzędnych G może (ale nie musi) być interpretowana jako czas.

Systemy typu (5) mogą być bezpośrednio powiązane z obiektami będącymi przedmiotem badań foresightowych. W tym celu zdefiniujemy scenariusze zdarzeń w systemie (5).

Definicja 1

Scenariuszem elementarnym s nazywamy sekwencję zdarzeń $(e_1, \dots, e_p)$, takich że jeśli $e_i = (q_i, q_{i+1})$, to $e_{i+1} = (q_{i+1}, q_{i+2})$.

Scenariusz elementarny jest zatem trajektorią systemu (5), określoną przez pewien ciąg decyzji i zdarzeń losowych $v_1, \dots, v_p$. Należy przy tym pamiętać, że dowolna sekwencja stanów nie musi być trajektorią systemu (5), gdyż poszczególne pary kolejnych stanów nie muszą tworzyć zdarzeń.

Zauważmy, że przez scenariusz ekspert foresightu rozumie zazwyczaj pewne uśrednione charakterystyki przyszłych procesów, zależne od określonych przyczyn. W badaniu foresightowym scenariuszy takich definiuje się z reguły od dwóch do pięciu, jest ich zatem znacznie mniej niż scenariuszy elementarnych w typowym systemie zdarzeń dyskretnych. Dla zachowania zgodności z powyższą, przyjętą powszechnie interpretacją scenariusza foresightowego, można przeprowadzić klasteryzację scenariuszy elementarnych, która w przedstawionym tu podejściu polega na zdefiniowaniu odległości pomiędzy scenariuszami elementarnymi, a następnie na tworzeniu klastrów, przy czym grupowanie odbywa się według dodatkowych zasad spójności grupy scenariuszy elementarnych. Tak utworzone scenariusze są zazwyczaj identyfikowane przez średnie wartości charakterystyk ilościowych ze zbioru $X(Q)$ w określonym zbiorze momentów w przyszłości.

Systemy zdarzeń dyskretnych mogą opisywać ewolucję systemu zdarzeń w otoczeniu PEEST instytucji, która tworzy lub wdraża strategię technologiczną. Model takiego systemu, a zwłaszcza związki przyczynowo-skutkowe i sprzężenia zwrotne, stanowi narzędzie do konstruowania scenariuszy elementarnych, które pojawiają się jako wyróżnione trajektorie systemu (5). W konsekwencji, używając aparatu metodologicznego systemów zdarzeń dyskretnych, do generowania niezdominowanych scenariuszy, strategii lub polityki można zastosować uogólnienie wielokryterialnego algorytmu wielokryterialnej najkrótszej ścieżki dla grafów o zmiennej strukturze, zależnej od sposobu dojścia do określonego wierzchołka (stanu). Idee rozwiązania tej klasy problemów pochodzą bezpośrednio z zasady programowania dynamicznego Bellmana [4; 6; 8]. Możemy użyć tej metody do uporządkowania stanów obiektu dla każdej chwili końcowej procesu modelowania, biorąc pod uwagę ilościowe charakterystyki wyznaczone przez algorytm najkrótszej ścieżki wielokryterialnej. Przypisując dodatkowe prawdopodobieństwa do każdego rozważanego zdarzenia zewnętrznego, możemy wyprowadzić oczekiwany ranking priory-

tetów dla każdej chwili z okresu modelowania. Wyniki mogą być przedstawione w postaci wykresu priorytetów jako funkcji czasu (przykład podany jest na rys. 2), przy założeniu wystąpienia w przyszłości określonych zdarzeń.

Główne etapy procedury wspomagania decyzji zwracającej wyniki zarówno ilościowe (oczekiwane przyszłe wartości wskaźników w obrębie każdego scenariusza), jak też i jakościowe (ranking priorytetów), są przedstawione niżej.

1. Nadrzędne cele polityczne, wskaźniki i dane makroekonomiczne są określone podczas fazy gromadzenia wiedzy.

2. Początkowe rankingi są konstruowane za pomocą metody analizy wielokryterialnej wybranej przez ekspertów wspomagających proces decyzyjny i przedyskutowane z osobami podejmującymi decyzje. W szczególności, do tworzenia rankingów rekomendowana jest metoda wielu punktów odniesienia, MREF, oraz metoda zbiorów odniesienia – REFSET 0, chociaż stosowane mogą być także inne metody rankingowe, w tym oparte o wyznaczenie współczynników liniowej agregacji kryteriów (tzw. scoring) [3; 9].

3. Po zdefiniowaniu obszarów polityki technologicznej $A_j, j=1, \dots, K$, do każdego z nich przyporządkowany może być początkowy Wskaźnik Relatywnej Ważności $\omega_j(t_0) \in (0, 1]$, na podstawie wytycznych zatwierdzonych przez władze. Wskaźnik ten umożliwia wstępne uporządkowanie obszarów interwencji opierając się na preferencjach politycznych (wyniki głosowań organów wybieralnych, badania opinii społecznej, preferencje władz nadrzędnych itp.).

4. Do opisu dynamiki rankingów i wyznaczenia rankingu docelowego konieczne jest uwzględnienie ewolucji wskaźników [12] i innych charakterystyk ilościowych rozważanych w badaniach foresightowych oraz analizy możliwości i tempa osiągnięcia celów strategicznych.

5. Ilościowe prognozy rozwoju zasobów ludzkich, wzrostu ekonomicznego, preferencji społecznych i przyszły budżet mogą być opracowywane przy zastosowaniu adaptacyjnej analizy trendu, na podstawie dostępnych danych statystycznych.

6. Należy scharakteryzować zbiór zdarzeń zewnętrznych, mogących wpłynąć na implementację projektów i ich dynamikę.

7. Należy scharakteryzować przyszłe scenariusze oraz ich związki ze wskaźnikami ilościowymi zdarzeń (zbiór $X(Q)$).

Konstrukcja rankingu, który jest częścią powyższej procedury może być zatem przedstawiona ogólnie jako:

ALGORYTM 1

Krok 1. Zdefiniuj obiekty, które mają być uporządkowane (ten krok może być częścią poprzednich etapów analizy i prognozowania).

Krok 2. Przypisz obiektom kryteria i określ ich wartości. Uwzględnij wstępne Wskaźniki Relatywnej Ważności obiektów, $\omega_j(t_0)$, jako jedno z kryteriów ilościowych.

Krok 3n. Zbierz i zaprezentuj wiedzę o przyszłości (trendy, scenariusze szacunki ekspertów, prognozy i ich wzajemne zależności) oraz jej wpływ na wartość kryteriów i innych elementów struktury preferencji decydentów.

Krok 4n. Zbuduj ranking w chwili t_0 do zastosowania na przedziale $[t_0, T)$ używając wybranej metody rankingowej.

Krok 5n. Przełóż rezultaty rankingów na rekomendacje ekspertów, jakie zostaną przedstawione osobom podejmującym decyzje.

Krok 6. Powtórz kroki 3n-5n po każdej zmianie decydentów, prognoz, parametrów lub okoliczności zewnętrznych.

3. Zastosowania metody DPR

Stosując przedstawioną wyżej procedurę i metody tworzenia rankingów opartej o zbiory odniesienia (REFSET [13]), możemy utworzyć dynamiczny ranking ważności priorytetów jako funkcji czasu.

Ogólna procedura opisana w algorytmie 1 może zostać uszczegółowiona poprzez dokładniejszy opis elementów dynamicznych naszkicowanych w poprzedniej sekcji. W rezultacie otrzymamy:

ALGORYTM 2 (DYNAMICZNE RANKINGI PRIORYTETÓW, DPR)

Krok 1. Zdefiniuj obiekty, które powinny być uszeregowane oraz opisz dynamikę ich ewolucji w okresie prognozowania rang.

Krok 2. Przypisz obiektom kryteria i określ ich wartości jako funkcje czasu. Dla początkowej chwili $n := t_0$ uwzględnij wstępne Wskaźniki Relatywnej Ważności obiektów, ω_j , $j = 1, \dots, K$, jako jedno z kryteriów ilościowych (jeśli nie są określone, wówczas można przyjąć $\omega_j = 1$ dla wszystkich j).

Krok 3. Zbierz i zaprezentuj wiedzę o przyszłości (trendy, zdarzenia, scenariusze, szacunki ekspertów, prognozy oraz ich wzajemne relacje oraz wpływ komponentów tej wiedzy na wartości kryteriów rankingowych, osiąganie celów strategicznych i innych elementów struktury preferencji decydentów (np. tzw. progi weta w metodzie ELECTRE Iv [1; 19])).

Krok 4n. Dla $n = t_0, \dots, N$ opisz ilościowo zależność rankingów priorytetów w chwili $(n+1)$ od rankingów w chwili n , uwzględniając zmiany wartości kryteriów, zmiany struktury preferencji będące rezultatem Kroku 3 oraz – dodatkowo – inercję rankingów, opisaną jako dodatkowe kryterium kosztowe. Jest ono proporcjonalne od odległości H rankingów w chwilach n i $(n+1)$ zdefiniowanej wzorem

$$H(r(n), r(n+1)) := \sum_{1 \leq k \leq K} s_k |r(n, k) - r(n+1, k)| / P(K) \quad (6)$$

gdzie $r(n, k)$ oznacza rangę k -tego obiektu w rankingu $r(n)$ w chwili n , a $P(K) = K(K-1)/2$ jest maksymalną odległością dwóch rankingów K obiektów. Współczynniki proporcjonalności $s_k > 0$ zależą od rzeczywistych kosztów zmiany rankingów poszczególnych obiektów (np. koszty publikacji ogłoszeń itp.). Mogą też zależeć od wartości $r(n, k) - r(n+1, k)$, różniąc się np. dla dodatnich i ujemnych wartości tej różnicy.

Krok 5n. Zbuduj ranking $r(n)$ dla chwili n za pomocą wybranej wielokryterialnej metody rankingowej (rekomendowaną metodą jest metoda zbiorów odniesienia REFSET). Następnie opierając się na $r(n)$ i rezultatach Kroku 4n dla $n+1$ utwórz ranking $r(n+1)$. Kontynuuj, aż osiągnięty zostanie horyzont planowania N .

Krok 6. Utwórz ranking dynamiczny $r(t_0, \dots, N)$. Wizualizuj zależność rang w $r(t_0, \dots, N)$ od czasu dla wybranych obiektów.

Ranking – pomimo znaczenia poznawczego – nie jest zazwyczaj ostatecznym celem analizy. Dlatego też rezultaty rankingów $r(t_0, \dots, N)$ powinny zostać przełożone na rekomendacje ekspertów, jakie zostaną przedstawione osobom podejmującym decyzje o przydziale środków finansowych na przedsięwzięcia powiązane z szeregowanymi technologiami lub projektami, biorąc pod uwagę także rezultaty innych metod analitycznych stosowanych w planowaniu oraz cele polityki osiągnięte wcześniej. Rekomendacje takie mogą mieć formę np. struktury preferencji określonej dla obiektów umiejscowionych w hierarchii polityki rozwojowej niżej od obszarów polityki lub nadrzędnych celów strategicznych (por. [13]), np. dla regionalnych projektów infrastrukturalnych. Przykład takiej procedury, gdzie algorytm 2 zastosowany został do alokacji środków na projekty, podany zostanie w następnej sekcji.

Otrzymany ranking $r(t_0, \dots, N)$ powinien być aktualizowany zgodnie ze zdefiniowanymi wcześniej zasadami polityki regionalnej za każdym razem, gdy osiągnięty zostanie jeden z celów strategicznych lub w przypadku istotnej zmiany warunków zewnętrznych.

Przed implementacją części obliczeniowej algorytmu 2, w celu uzyskania adekwatnych procedur wspomagania decyzji należy zbudować model struktury informacyjnej opisanej w Kroku 4 algorytmu 2. W szczególności, należy odpowiednio zinterpretować zależności ilościowe oraz uwzględnić wpływ zmian technologicznych i wyników badawczych na atrybuty charakteryzujące populację lub badany zbiór obiektów [12] w modelu optymalizacji wielokryterialnej.

Ważnym elementem procedury wspomagania decyzji jest odpowiednia wizualizacja otrzymanych rezultatów w postaci wykresu rang priorytetyzowanych obiektów w funkcji czasu. Wykres taki może zostać rozszerzony o wskaźniki pokazujące, również jako funkcje czasu dla tego samego okresu, stopień społecznego i ekonomicznego wpływu na osiągnięcie szeregowanych celów polityki technologicznej.

Dla zilustrowania idei dynamicznych rankingów przedstawionych w niniejszym rozdziale mniej istotny jest wybór metody rankingowej stosowanej w każdym kroku algorytmu 2. W związku z tym, dla uproszczenia, w poniższym przykładzie założymy, że statyczne rankingi tworzone w każdym kroku procedury dynamicznej oparte są na najprostszej metodzie scoringowej (przegląd takich metod podany jest w [13], chociaż ogólne zasady postępowania podane w algorytmie 2 pozwalają stosować także metody niekompensacyjne typu Electre czy Promethee [2]. Najprościej zastosować kryterium scoringowe o interpretacji „dobrobyt ogółu mieszkańców regionu” i założyć istnienie pewnych wag umożliwiających utworzenie sumy ważonej ze wskaźników charakterystycznych dla każdego z „obszarów” polityki. W przykładzie 1 zastosujemy ponadto dalsze uproszczenia:

- koszt „inercji” rankingu nie powinien być zbyt wysoki, z każdym priorytetyzowanym obszarem polityki regionalnej powiązany jest jeden cel reprezentowany przez jeden wskaźnik ilościowy,
- w każdej chwili dostępny jest tylko jeden scenariusz,
- wykorzystywane przy aktualizacji rankingów prognozy dotyczą jedynie czasów i stopnia osiągnięcia celów związanych z każdym z priorytetyzowanych obszarów polityki regionalnej.

Przykład 1

Metody obliczeniowe i wizualizacja rezultatów priorytetyzacji pro-rozwojowych polityk technologicznej w jednym z regionów.

A. Priorytetyzowane obszary polityki, cele i wskaźniki

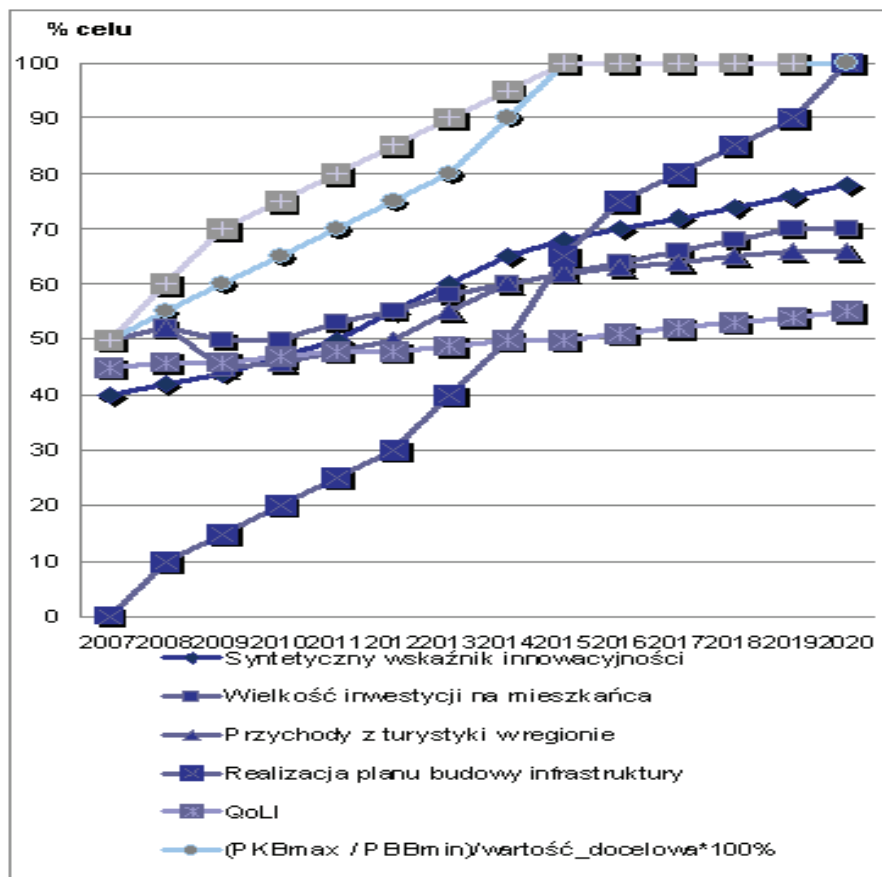
Tabela 1

Przykład celów politycznych mających wpływ na ocenę technologii

Lp.	Obszar polityki regionalnej	Cel polityki/rodzaj celu	Syntetyczny wskaźnik ilościowy
1	Warunki dla rozwoju społeczeństwa opartego na wiedzy	Osiągnięcie wskaźników charakteryzujących najbardziej rozwinięte regiony w UE (wielowskaźnikowy)	Suma ważona ilości patentów, sprzedaży produktów zaawansowanych technologii, ilości doktoratów itp.
2	Gospodarka regionalnej szansy	Trajektoria wartości inwestycji, w tym w sektorze high-tech	Skumulowana wartość inwestycji na mieszkańca pomniejszona o ich amortyzację
3	Turystyka i kultura	Trwały wzrost wpływów z turystyki, osiągnięcie określonego opisowo poziomu usług	Skumulowana suma wpływów z turystyki w regionie w okresie planowania
4	Infrastruktura dla rozwoju gospodarczego	Realizacja szczegółowego planu budowy infrastruktury komunikacyjnej i telekomunikacyjnej	Procentowy (rzeczowy) stopień realizacji planu budowy infrastruktury (tele)komunikacyjnej
5	Rozwój Obszaru Metropolitalnego	Zwiększenie atrakcyjności centrum metropolitalnego	Wskaźnik jakości życia (QoLI według the Economist)
6	Spójność wewnątrzregionalna	Zmniejszenie różnicy poziomu życia i PNB na mieszkańca pomiędzy gminami regionu do co najwyżej 50%	Stosunek PNB na mieszkańca w wieku produkcyjnym pomiędzy gminami regionu
7	Infrastruktura ochrony środowiska	Osiągnięcie docelowych wartości wskaźników charakteryzujących czystość wód, powietrza, udział recyklingu i odpadów sortowanych itp.	Regionalny indeks oddziaływania na środowisko (suma ważona indeksów cząstkowych dla rodzajów oddziaływań i jednostek administracyjnych)

B. Założenia

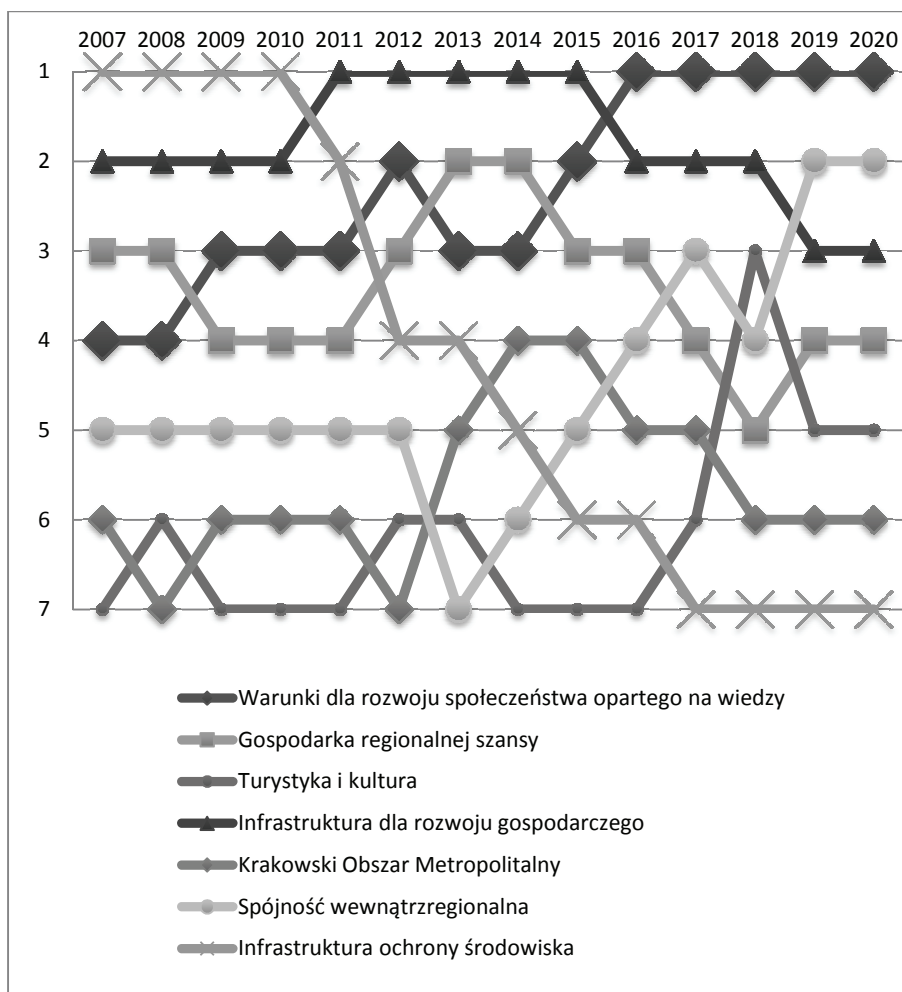
1. Dostępne prognozy: dla każdego celu związanego z obszarem polityki zakładamy znajomość trajektorii jego osiągnięcia, tj. wartości powiązanego syntetycznego wskaźnika jako funkcji czasu (rys. 1).



Rys. 1. Założone wartości wskaźników osiągnięcia celów regionalnych zastosowane w metodzie DPR dla analizy priorytetów Małopolski w okresie 2007-2020 (w % wyznaczonego celu)

2. Chwilowe wartości współczynników „ważności obszaru polityki” $\omega_j(t)$, $0 < \omega_j(t) \leq 1$, $j = 1, \dots, 7$, $2007 \leq t \leq 2020$, zależą liniowo od procentowego stopnia zostało osiągnięcia odpowiedniego wskaźnika w danej chwili (im więcej, tym są mniejsze) oraz od wartości początkowych dla $t = 2007$.

3. Wartości wszystkich wskaźników ilościowych są znormalizowane i należą do przedziału $[0; 1]$.



Rys. 2. Przykład zastosowania metody DPR dla analizy priorytetów w okresie 2007-2020

4. Funkcja scoringowa, służąca do porządkowania obszarów polityki regionalnej dla każdego t , ma postać sumy poszczególnych wskaźników ilościowych ważonej przez współczynniki $\omega_j(t)$,

5. Inercja rankingów obliczana jest według wzoru (6), przy czym wszystkie wartości $s_k(t)$ są równe 0,01.

Wykres dynamicznych rang priorytetów obszarów polityki regionalnej, będących wynikiem zastosowania algorytmu 2 do powyższej struktury danych przedstawiono na rys. 2.

4. Zastosowanie rezultatów priorytetyzacji technologii do alokacji środków

Wyniki algorytmu 2 oraz opisane wyżej podejście dynamicznego rankingu mogą być zastosowane w następującej procedurze alokacji zasobów na projekty technologiczne, która wykorzystuje systemy zdarzeń dyskretnych do modelowania niezależnych od decydenta zjawisk i procesów wpływających na alokację:

ALGORYTM 3

Struktura danych

1. Charakterystyka stanu obecnego regionu, gdzie wdrażana jest strategia – w postaci wektora $X(Q)$, którego współrzędne są wartościami odpowiednich wskaźników.

2. Ograniczenia na wysokość alokacji środków, łącznie, w odniesieniu do branż i podregionów oraz w każdym okresie planowania.

3. Zbiór $S(0)$ scenariuszy początkowych przyszłych stanów systemu rozpoczynających się w ustalonym podzbiorze Q_0 stanów początkowych systemu Q w chwili t_0 , będącej zakładanym początkiem wdrażania strategii (gdy chwila t_0 jest bliska chwili tworzenia strategii, wówczas najczęściej $Q_0 := \{q\}$). Scenariusze i prognozy są rezultatem przeprowadzonych wcześniej badań foresightowych [15; 18].

4. Algorytm η tworzenia rankingu projektów r w zależności od rankingu priorytetowych celów strategicznych $R := (C_1, \dots, C_l)$. Zakładamy przy tym, że realizacja każdego projektu p jest działaniem zmieniającym stan systemu Q w ten sposób, że możliwe jest ilościowe określenie $\varphi_i := \varphi(C_i)$ (np. jako wskaźnik procentowy) stopnia osiągnięcia każdego z celów priorytetowych C_i po zakończeniu projektu. Zakładamy ponadto, że dla każdego celu strategicznego C_i z rankingu R określona jest funkcja $\alpha_i(t)$ interpretowana jako (dodatni) współczynnik istotności celu C_i i spełniająca warunek

$$C_i \triangleleft_t C_j \Rightarrow \alpha_i(t) < \alpha_j(t)$$

gdzie „ $\triangleleft_t$ ” jest porządkiem w rankingu celów strategicznych zależnym od czasu. Reguła decyzyjna algorytmu η może być wówczas zdefiniowana przez warunek

$$p_k \leq_t p_l \Leftrightarrow \sum_{1 \leq i \leq l} \alpha_i(t) [\varphi_i(p_k) - \varphi_i(p_l)] < 0$$

z wynikowym porządkiem „ $\leq_t$ ” stosowanym w rankingu projektów r .

5. Reguła alokacji funduszy na projekty γ w zależności od rankingu projektów r . Obowiązuje ogólna zasada, że środki otrzymują projekty najwyżej ocenione, jednak γ musi uwzględniać zarówno ograniczenia budżetowe w każdym z okresów, jak i ograniczenie na łączną wysokość alokowanych środków. Ponadto zakładamy, że rezygnacja z projektu, będąca wynikiem zmiany rankingu, związana jest z pewną stratą zależną od stanu realizacji projektu (brak opcji rezygnacji) oraz że strata ta może być agregowana z jednym z kryteriów oceny projektów.

6. Kryterium ewaluacyjne G , określające jakość alokacji środków w całym okresie planowania.

Krok 1. Za pomocą wybranej metody analizy wielokryterialnej znajdź rankingi (w funkcji scenariusza)

$$r(t_0; S(0)) := \{(p_{1,s(0)}, \dots, p_{k(0),s(0)}) : s(0) \in S(0)\}$$

dla projektów, które mają być sfinansowane w chwili $t = t_0$.

Krok 2. Zastosuj regułę alokacji funduszy $\gamma := \gamma(p, s)$ dla projektów $(p_{1,s(t)}, \dots, p_{k(0),s(t)})$ i scenariusza s , gdzie $\gamma(p, s)$ jest macierzą o wymiarze $k(0) \times T$.

Krok 3n. Wyznacz wartość kryterium jakości alokacji środków $G(0) := G(0, \gamma)$. Zastosuj regułę sterowania δ do zdarzeń tworzących scenariusze ze zbioru $S(n)$ i do ewentualnych zdarzeń zewnętrznych.

Krok 4n. Aktualizuj rankingi (w funkcji scenariusza) $r(t_0, S(0)) = (p_{1,s(t)}, \dots, p_{k(0),s(t)})$ projektów, które mają być sfinansowane dla $t = t_0$ określone w Kroku 1. Znajdź wartość kryterium jakości alokacji środków $G(n)$. Jeśli $G(n) \leq G(n-1)$, zbadaj zależność między wartością $G(n) - c(n, S(n)) + g(n, S(n))$ a $G(n-1)$, wyznacz wartość optymalną. Zastosuj sekwencję operacji $(v(1,1), \dots, v(1,k))$, $v(i,j) \in V$ zgodnie z (5), odpowiadającą wybranemu scenariuszowi aktualizacji, wyznacz nową wartość $G'(n)$.

Krok 5n. Zastosuj regułę alokacji środków γ do $(p_{1,n(t)}, \dots, p_{k(n),n(t)})(s)$ dla scenariusza aktualizacji s .

Krok 6n. Powtórz kroki 3n-5n dla n zmieniającego się w okresie alokacji środków.

Krok 7. Wyznacz wartości końcowe kryteriów alokacji zasobów.

Scenariusze zastosowane w powyższym schemacie przydziału środków mogą być generowane w sposób typowy dla klasycznych metod foresightu [15; 18]. Mogą też być rezultatem zastosowania metod generowania najpierw scenariuszy elementarnych w systemach zdarzeń dyskretnych i następnie ich klasteryzacji. Niezależnie od metody, wymagane jest by opis scenariusza był spójny z opisem systemu (5), $X(Q)$ oraz z parametrami celów strategicznych C_i . Możliwe jest wówczas zastosowanie metod prognozowania wskaźników systemu „wewnątrz” każdego scenariusza, tj. tworzenie prognoz warunkowych,

przy założeniu zajścia zdarzeń charakteryzujących dany scenariusz. Ze względu na zazwyczaj stosunkowo niewielką liczbę danych w powstałych w ten sposób szeregach czasowych, celowe jest ich wstępne wygładzenie (czyli ograniczenie wariancji) poprzez transformacje uśredniające lub agregację obwiedni ekstremalnych.

Podsumowanie

Opisany sposób postępowania powinien zapewnić lepsze dopasowanie procedur rozdziału funduszy przeznaczonych na inwestycje do zmieniających się potrzeb i okoliczności zewnętrznych. Konkretna implementacja algorytmów, w szczególności wybór metody rankingowej stosowanej w celu konstrukcji indywidualnych rankingów w każdej chwili okresu prognozowania, musi być dostosowany do potrzeb decydentów i specyfiki interesariuszy rankingu. Warto zauważyć, że pomimo istnienia ogólnych zasad tworzenia wieloletnich strategii inwestycyjnych, w konkretnych zastosowaniach indywidualnemu doborowi podlegają metody tworzenia i charakterystyki scenariuszy, metody rankingowe, reguły uzależniające rankingi projektów technologicznych od uszeregowania priorytetów wynikających z polityki (regionalnej, krajowej lub unijnej), a także operacyjne reguły alokacji środków na projekty. W związku z tym, w celu efektywnego zastosowania przedstawionych tu podstaw dynamicznej adaptacji strategii inwestycyjnych i opartych na nich procedur do zmiennych warunków otoczenia wskazane jest, by przed rozpoczęciem opracowywania dokumentów strategicznych decydenci nawiązali współpracę z instytucjami naukowymi zajmującymi się metodami analizy wielokryterialnej.

Oprócz problemów związanych z rozdziałem środków publicznych, przedstawiona tu metodyka może być stosowana do wyboru inwestycji finansowych [14], w szczególności do uszeregowania akcji spółek wybieranych do portfela biorąc pod uwagę scenariusze rozwoju sytuacji na rynku kapitałowym [17]. Podobne metody stosowane są także do oceny innowacji technologicznych zgłaszanych na konkursy organizowane przez Centra Transferu Technologii [16]. Należy przy tym mieć na uwadze, że dla każdej klasy zastosowań powinna być zaprojektowana indywidualnie dostosowana architektura systemu wspomagania decyzji wielokryterialnych z odpowiednio zintegrowaną bazą wiedzy. Taki system wspomagania decyzji może być zastosowany również do analizy rezultatów foresightu branżowego czy regionalnego i – w razie potrzeby – do aktualizacji celów polityki sektorowej czy regionalnej, a także celów korporacyjnej strategii inwestycyjnej.

Literatura

1. Barrios S., Strobl E. (2009). The Dynamics of Regional Inequalities. *Regional Science and Urban Economics*, 39, s. 575-591.
2. Bouyssou D., Marchant T. (2007). An Axiomatic Approach to Noncompensatory Sorting Methods in MCDM. The Case of Two Categories. *European Journal of Operational Research*, 178, s. 217-245.
3. Chebotarev P.Yu., Shamis E. (1998). Characterizations of Scoring Methods for Preference Aggregation. *Ann. Oper. Res.*, 80, s. 299-332.
4. Henig M.I. (1985). The Shortest Path Problem with Two Objective Functions. *European Journal of Operational Research*, 25, s. 281-291.
5. Kaliszewski I. (2008). Wielokryterialne podejmowanie decyzji. Obliczenia miękkie dla złożonych problemów decyzyjnych. WNT, Warszawa, s. 154.
6. Martins E.Q.V. (1984). On a Multicriteria Shortest Path Problem. *European Journal of Operational Research*, 16, s. 236-245.
7. Medaglia A.L., Hueth D., Mendieta J.C., Sefair J.A. (2008). A Multiobjective Model for the Selection and Timing of Public Enterprise Projects. *Socio-economic Planning Sciences*, 42, s. 31-45.
8. Opasanon S., Miller-Hooks E. (2006). Multicriteria Adaptive Paths in Stochastic, Time-varying Networks. *European Journal of Operational Research*, 173, s. 72-91.
9. Saaty T.L. (2006). Rank from Comparisons and from Ratings in the Analytic Hierarchy/Network Processes. *European Journal of Operational Research*, 168, s. 557-570.
10. Salo A., Mild P., Pentikäinen T. (2006). Exploring Causal Relationships in an Innovation Program with Robust Portfolio Modeling. *Technological Forecasting and Social Change*, 73, s. 1028-1044.
11. Skulimowski A.M.J. (1991). Optimal Control of a Class of Asynchronous Discrete-Event Systems. W: *Automatic Control in the Service of Mankind. Proceedings of the 11th IFAC World Congress, Tallinn (Estonia), Aug. 1990*. Pergamon Press, London, 3, s. 489-495.
12. Skulimowski A.M.J. (1994). Optimizing the Structure of a Partitioned Population. W: *System Modelling and Optimization; Lecture Notes in Control and Information Sciences*. Red. J. Henry i J.P. Yvon. Springer-Verlag, Berlin-Heidelberg-New York, 197, s. 771-782.
13. Skulimowski A.M.J. (1997). Methods of Multicriteria Decision Support Based on Reference Sets. W: *Advances in Multiple Objective and Goal Programming, Lecture Notes in Economics and Mathematical Systems*. Red. R. Caballero, F. Ruiz, R.E. Steuer. Springer-Verlag, Berlin-Heidelberg-New York, 455, s. 282-290.

14. Financial Modelling – Proceedings of the 23rd Meeting of the EURO WG on Financial Modelling. Red. A.M.J. Skulimowski. (1999). Cracow, October 8-10, 1998. Progress & Business Publishers, Kraków, s. 504.
15. Skulimowski A.M.J. (2006). Framing New Member States and Candidate Countries Information Society Insights. W: Prospects for a Knowledge-Based Society in the New Members States and Candidate Countries. Red. R. Compano, C. Pascu. Publishing House of the Romanian Academy, s. 9-51.
16. Transfer Technologii w Informatyce i Automatyce (Technology Transfer in Computer Science and Automation). (2006). Red. A.M.J. Skulimowski. Progress & Business Publishers, Kraków, 2006, s. 406.
17. Skulimowski A.M.J. (2008). Application of Dynamic Rankings to Portfolio Selection. W: New Developments in Financial Modelling. Red. J.O. Soares, J.P. Pina, M. Catalão-Lopes. Newcastle, CSP Cambridge Scholars Publishing, 2008. Proceedings of the 41st Meeting of the Euro Working Group on Financial Modelling. Lisbon, Portugal, November 8-9, 2007, s. 196-212.
18. Priorytetyzacja celów polityki regionalnej w kontekście efektywnego wykorzystania Funduszy Strukturalnych. Raport Końcowy projektu nr POPT.03.01.00-00-002/07zrealizowanego w roku 2008 w ramach I edycji Konkursu Dotacji Fundusze Europejskie na poziomie NSS. (2009). Red. A.M.J. Skulimowski. Kraków.
19. Trzaskalik T. (2011). Metody rankingowe w wielokryterialnym podejmowaniu decyzji. W: Metody wielokryterialnego wyboru i konstrukcji rankingów projektów finansowanych ze środków publicznych. Red. A.M.J. Skulimowski. Wydawnictwo Progress and Business, Kraków, s. 31-55.
20. Yen-Liang Chen, Li-Chen Cheng (2010). An Approach to Group Ranking Decisions in a Dynamic Environment. Decision Support Systems, 48, s. 622-634.

METHODS OF FORECASTING DYNAMIC RANKINGS OF TECHNOLOGICAL PRIORITIES

Summary

This paper is devoted to the elaboration of methodological foundations of forecasting models that may be applied to establish anticipated future rankings of key technologies or strategic goals. We will also present the resulting practical approaches for rankings arising in some relevant classes of multicriteria decision problems. Usually, such rankings are built in within strategic planning activities and rely on results of prior as well as simultaneous foresight exercises. The rankings of technological priorities can then be applied to allocate the organization's budget to different investment areas covered by the competences of the authorities concerned. Another common approach to apply the rankings is to define the temporal relevance of priorities, i.e. those ranked higher are considered first. We will show how the dynamic rankings can be constructed from forecasts or foresight results concerning the external circumstances, economical,

social, and technological, using a dynamic programming model. The forecasting model presented in Secs. 3-5 contains the rules which govern the actions of the decision makers when a technological priority changes in a response to an external event or as a result of achieving certain goals. The novelty of the *dynamic priority ranking approach* (DPR) here presented consists a.o. in the fact that the mutual impacts of events forming and influencing future scenarios are taken into account in one model jointly. The forecasting results are visualised as a time-dependent relative priority importance chart showing the priorities on an ordinal scale. In Sec.5 we will provide an illustrative example of the real-life applications of the adaptive ranking methodology to rank key intervention areas in a foresight exercises, where the ranking forecasts have been derived using the reference-set-based outranking method. Using an assignment algorithm provided in Sec. 5, the results can be applied to assign funding to innovative technological projects.

Jacek Szołtysek

Grażyna Trzpiot

Uniwersytet Ekonomiczny w Katowicach

DRZEWIA KLASYFIKACYJNE W BADANIU PREFERENCJI KOMUNIKACYJNYCH

Wprowadzenie

Zachowania komunikacyjne to decyzje w zakresie sposobu podróży (pokonywania przestrzeni) oraz sam proces realizacji przemieszczenia. Decyzje o sposobie podróżowania podejmowane są na podstawie poglądów, opinii, przekonań, kształtujących wyobrażenie o tym, w jaki sposób można najlepiej dla siebie pokonywać przestrzeń. Ta uwaga jest niezwykle istotna w aspekcie prób zmiany zachowań komunikacyjnych w miastach, które wymagają, by dla dobra ogółu (a zatem i w pewnym zakresie również osoby podejmującej decyzje o podróży) zrezygnować z części wygod i zdecydować się na mniej atrakcyjny sposób pokonywania przestrzeni.

Poglądy, opinie i przekonania stanowią pewien „światopogląd” służący podejmowaniu decyzji komunikacyjnych. Przebieg i wynik procesów decyzyjnych, będących składową zachowań komunikacyjnych, nie różni się od wszelkich innych procesów podejmowania decyzji przez nabywców – stąd działania i czynności, jakie podejmuje konsument zależą od jego postawy, która ukierunkowuje działania jednostki i jest źródłem określonego sposobu zachowania. Nierzadko istotną rolę odgrywają dane zawarte w pamięci jednostki, inne ściśle związane ze spostrzeganiem procesy poznawcze, takie jak uwaga, pamięć, uczenie się, procesy emocjonalno-motywacyjne, struktura i funkcjonowanie osobowości jednostki. Ważną cechą psychiczności człowieka jest tzw. obszar jego prywatności. Oznacza to, że każdy z nas posiada potrzebę izolacji od otoczenia. Cecha ta, z punktu widzenia zachowań komunikacyjnych w mieście, przekształca się w tzw. egoizm kooperacyjny, zakładający konieczność współpracy z otoczeniem, jednak wyznacza jej granice. Zadaniem kształtowania zachowań komunikacyjnych w mieście jest zwiększanie zakresu egoizmu kooperacyjnego.

Dlaczego należy kształtować zachowania komunikacyjne w mieście? Głównym powodem jest chęć, a raczej konieczność zmniejszania zatłoczenia komunikacyjnego współczesnych miast. Skutek nagromadzenia dużej ilości przemieszczających się w stosunkowo krótkim czasie na niewielkiej przestrzeni miasta, przeznaczonej do pokonywania przestrzeni (infrastruktura liniowa i punktowa transportu), jak też pretendowaniu do jej zajęcia przez pojazdy wożące ładunki, powoduje zatłoczenie komunikacyjne, zwane kongestią. To zjawisko, zwane popularnie „korkami”, wywołuje wiele negatywnych zdarzeń (skutków), zarówno w wymiarze miasta, jak też poszczególnych grup jego użytkowników. W zasadzie nie ma takiego podmiotu lub osoby, dla której zatłoczenie komunikacyjne w dłuższej perspektywie jest korzystne. Dla miasta, czy też regionu, długotrwale utrzymujące się zatłoczenie komunikacyjne jest bardzo istotną przeszkodą w rozwoju i skutecznym konkurowaniu z innymi miastami czy regionami. Taka sytuacja, jak wynika z badań, prowadzi do upadku miast czy regionów.

Podjmowanie decyzji o sposobie podróżowania w przejazdach obligatoryjnych znacząco odbiega od procesu decyzyjnego w przewozach fakultatywnych (rys. 1). O ile w przypadku podróży okazjonalnych rozważane są jej warianty i dokonywany jest wybór jak najbliższy zestawowi wymagań w zakresie czynników waloryzacji takiej podróży, to w przypadku podróży obligatoryjnych mamy do czynienia z uproszczoną i – niestety – utrwaloną procedurą, składającą się na także utrwalone zachowania komunikacyjne. Proces postrzegania jest złożony, aktywny, cykliczny, ma charakter decyzyjny oraz w dużej mierze nieświadomy. Percepcja umożliwia kształtowanie stałego, zorganizowanego, spójnego i sensownego odzwierciedlenia rzeczywistości. Obligatoryjny (bezwzględnie lub względnie obligatoryjny) charakter mają te potrzeby przewozowe, których realizacja warunkuje egzystencję człowieka. Potrzeby te występują systematycznie i powodują konieczność codziennych przemieszczeń w dniach pracy i zajęć w ściśle określonych godzinach i w tych samych relacjach (potrzeby bezwzględne). Potrzeby względnie obligatoryjne wymagają również stałych przemieszczeń, lecz niekoniecznie codziennych, podejmowanych w dowolnych godzinach w niestałych relacjach. Proces postrzegania jest tu praktycznie wyłączony przez rutynę.



Rys. 1. Proces decyzyjny w zakresie wyboru sposobu realizacji potrzeby komunikacyjnej

Porównując cechy przejazdów obligatoryjnych i fakultatywnych należy w świetle niniejszych rozważań zwrócić uwagę na dwa aspekty: ich częstotliwość w połączeniu z istotnością życiową oraz uproszczony sposób podejmowania decyzji (jak w przypadku zakupów codziennego użytku – rutynowych^{*}). Raz błędnie (lub niekorzystnie) podjęta decyzja o sposobie przemieszczania jest powielana długo i bez należytej refleksji. Przyzwyczajenie to nic innego jak nieświadome trwanie w pewnych utrwalconych już nawykach. Złe nawyki powielane codziennie w podróży obligatoryjnych przez ogromną ilość podróżujących to poważny problem dla miasta czy regionu. Stąd powstała koncepcja kształtowania odmiennych, korzystnych dla miasta czy regionu, zachowań komunikacyjnych. Te zmiany mają doprowadzić do pożądanego sposobu pokonywania przestrzeni miejskiej. Chodzi o to, by unikać nagromadzenia dużej ilości osób pretendujących do korzystania w tym samym czasie z infrastruktury transportowej miasta w warunkach, gdy jej zdolność przepustowa zbliża się do wyczerpania. Zmiany zachowań polegają na takim wyborze środków transportu w podróży, który preferuje komunikację zbiorową i marginalizuje komuni-

^{*} Rutynowe zakupy to jednak nie tylko te codzienne zakupy, ale także te, które robimy co jakiś czas, a które zawsze są konieczne i nieuniknione.

kację indywidualną, co wymaga od użytkowników samoograniczania i wyrzeczeń. Zazwyczaj chodzi o zmniejszenie komfortu podróży bądź elastyczności połączeń, czasu podróży i innych czynników waloryzacji usługi transportowej. Chodzi zatem o to, by znając preferencje komunikacyjne ludności (czy raczej – jej oddzielnych segmentów) znaleźć sposoby (organizacyjne, ideologiczne czy inne) zmiany postrzegania* nowego i dotychczasowego sposobu przemieszczania się. We współczesnej psychologii istnieją dwa podstawowe podejścia w analizie procesu spostrzegania. Pierwsze z nich dotyczy wyodrębniania cech. Według niego istotą percepcji jest integrowanie prostych elementów bodźca w jedną, złożoną całość. Przedmiot, który jest rozpoznawany zostaje porównany do określonego wzorca (poznawczej reprezentacji tego przedmiotu w pamięci). Porównywanie do wzorca odbywa się najczęściej w warunkach korzystnych dla jednostki. Druga z koncepcji podkreśla tworzenie bądź odkrywanie przedmiotu w procesie organizacji percepcyjnej. Proces ten łączy nastawienie, oczekiwanie i posiadane doświadczenia, które są zakodowane w strukturach poznawczych człowieka. Wiedza, procesy pamięci, wnioskowania stanowią podstawę spostrzegania. Człowiek odbierając różnorodne informacje stara się łączyć je i organizować tak, aby dostrzec ich znaczenie i sens**. Kształtowanie zachowań komunikacyjnych bazuje ponadto prawdopodobnie na koncepcji braku***. Brak ma wartość ujemną, podczas gdy posiadanie ma wartość dodatnią. Potrzeba wyraża brak, zaś nie uświadamianie braku – brak potrzeby. Możliwe jest tworzenie braków przez ich uświadamianie. Uświadamiamy braki, by je likwidować. Do tego celu używamy narzędzi kształtujących zachowania komunikacyjne – w głównej mierze bazujących na perswazji. W działaniu perswazyjnym wykorzystuje się trzy techniki oddziaływania na odbiorcę: apelowanie o zajęcie określonego stanowiska i podjęcie określonego działania, sugerowanie pożądanых interpretacji i ocen oraz racjonalne uzasadnianie słuszności prezentowanych poglądów [1]. Wszystkie opisane mechanizmy zastosowane w praktyce miasta czy regionów mogą być przesłanką sukcesu jedynie, gdy bazują na dobrze zidentyfikowanych preferencjach komunikacyjnych.

* Spostrzeganiem nazywa się złożony układ procesów prowadzący do ukształtowania się subiektywnego obrazu rzeczywistości, spostrzeżenia

** <http://psychologia.net.pl/slownik.php?level=30> (24.03.2011).

*** Piszemy „prawdopodobnie”, gdyż jest to nasze logiczne spostrzeżenie, nie potwierdzone jednak żadnymi badaniami źródłowymi.

1. Drzewa decyzyjne w badaniu preferencji

Wiele zmiennych w badaniach prowadzonych na gruncie ekonomii jest obserwowanych na słabych skalach (nominalnej i porządkowej). W pracy proponujemy wykorzystanie do opisu preferencji metody klasyfikacji udzielonych odpowiedzi. Klasyfikację zbioru zmiennych przeprowadzimy wykorzystując nieparametryczne metody klasyfikacji – drzewa klasyfikacyjne.

Metody te dokonują w sposób rekurencyjny podziału m -wymiarowej przestrzeni X^m w jednorodne podzbiory ze względu na zmienną zależną y . W badaniach wykorzystano dwa rodzaje drzew klasyfikacyjnych: typu C&RT oraz CHAID. Drzewo klasyfikacyjne można zapisać następującym równaniem [2]

$$y = \sum_{k=1}^K a_k I\{\mathbf{x}_i \in R_k\} \quad (1)$$

gdzie:

$\mathbf{x}_i$ – wielowymiarowa obserwacja, klasyfikowane obiekty, element zbioru X^m ,
 $R_k - k = 1, \dots, K$ – rozłączne zbiory (segmenty) m -wymiarowej przestrzeni X^m ,
 a_k – zbiór parametrów,

$I(q)$ – funkcja wskaźnikowa: $I(q) = \begin{cases} 1 & , q \text{ jest prawdziwe} \\ 0 & , q \text{ jest fałszywe.} \end{cases}$

Na efektywność wyników klasyfikacji otrzymanych za pomocą drzewa C&RT wpływa między innymi dobór zmiennych. Do oceny efektywności klasyfikacji wykorzystuje się miary jakości podziału, które badają stopień zróżnicowania pomiędzy otrzymanymi podzbiorymi zmiennych ze względu na wartości rozpatrywanej w danej klasyfikacji zmiennej zależnej. Wśród wielu miar stosowanych w tym celu w badaniach wykorzystano wskaźnik zróżnicowania Giniego. Wykorzystana miara Gini'ego niespójności węzła jest często używana w sytuacjach, gdy zmienna zależna jest zmienną skategoryzowaną. Efektywność klasyfikacji nie jest tylko związana z doбором zmiennych. W wyniku podziału wielowymiarowego zbioru zmiennych, chcielibyśmy również otrzymać czytelne w interpretacji drzewo. Wymaga to zatrzymania rekurencyjnego podziału. Rezultat podziału, w którym otrzymujemy jednoelementowe zbiory danych nie jest interesujący z punktu widzenia klasyfikacji. Problem ten wy-

maga odpowiedniego kryterium przerywania podziału. Wykorzystano regułę bezpośredniego zatrzymania FACT (Fast Algorithm for Classification Trees) dla zadanej frakcji obiektów.

Poza miarą dokładności, w zagadnieniach klasyfikacyjnych oblicza się miarę niejednorodności węzła miarę Chi-kwadrat (χ^2). Zapiszemy poniżej wykorzystane miary:

- współczynnik Gini'ego [2]

$$H(S_k) = 1 - \sum_{j=1}^J p^2(j/k) \quad (2)$$

- statystyka χ^2

$$\chi^2 = \sum_{j=1}^J \sum_{k=1}^K \frac{(N_j(k) - E_j(k))^2}{E_j(k)} \quad (3)$$

gdzie:

J – liczba obserwacji zmiennej zależnej y ($j = 1, \dots, J$),

K – liczba rozłącznych podzbiorów zbioru S ($S = S_1 \cup S_2 \cup \dots \cup S_K$, $k = 1, \dots, K$),

N – liczba obserwacji w próbie ($N = \sum_{j=1}^J N_j = \sum_{k=1}^K N_j(k)$),

$E_j(k)$ – oczekiwana liczba y równa j w k -tym podzbiorze S
 $(E_j(k) = \frac{N_j \cdot N(k)}{N})$,

$p(j/k)$ – prawdopodobieństwo zdarzenia zmienna zależna y wynosi j w segmencie R_k , gdzie j jest najczęstszą wartością zmiennej y .

Zagadnienia omówione we odniesieniu do drzew C&RT odnosi się również do CHAID. Obydwie metody tworzą drzewa, w których każdy węzeł (oprócz liści) zawiera warunek podziału, a celem jest klasyfikacja (jakościowej zmiennej zależnej). Akronim CHAID oznacza Chi-squared Automatic Interaction Detector. Jest to jedna z pierwszych metod drzew klasyfikacyjnych, zaproponowana przez Kassa [3]. CHAID buduje drzewa, w których z węzłów mogą wychodzić więcej niż dwie gałęzie. Podstawowy algorytm używany do budowania drzew (nie-binarnych), w problemach klasyfikacyjnych (gdzie zmienna zależna jest ze swej natury jakościowa) opiera się na teście χ^2 jako kryterium wyznaczenia kolejnego najlepszego podziału w każdym kroku.

Wyznaczany jest taki podział, który daje największą istotność testu χ^2 wyznaczanego dla tablicy kontyngencji. Algorytm CHAID można opisać następująco: na wstępie tworzymy predyktory jakościowe z ilościowych poprzez podział rozkładu na pewną ilość kategorii tak, by do każdej wpadało w przybliżeniu tyle samo obserwacji. W przypadku predyktorów jakościowych klasy są już zdefiniowane w sposób „naturalny”.

W kolejnym kroku przeglądamy po kolei predyktory i wyszukujemy dla każdego z nich pary kategorii, które różnią się najmniej pod względem zmiennej zależnej. W przypadku problemów klasyfikacyjnych (gdzie zmienna zależna jest też jakościowa) ocena wykonywana jest przez obliczenie testu χ^2 . Jeśli test dla danej pary kategorii nie daje różnicy istotnej statystycznie na *poziomie p dla łączenia*, program łączy te kategorie i powtarza ten krok (tzn. szuka kolejnej pary kategorii, w której może znaleźć się też kategoria dopiero co powstała). Jeśli wartość p jest istotna statystycznie (mniejsza niż odpowiednia wartość *poziomu p dla łączenia*), obliczamy (opcjonalnie) poprawkę Bonferroniego* p -value dla zbioru kategorii predyktora.

W kolejnym kroku wybieramy predyktor o najniższej wartości *poziomu p* (z poprawką), tzn. ten predyktor, który daje najbardziej istotny podział. Jeśli wartość *poziomu p* (z poprawką Bonferroniego) dla każdego predyktora jest niższa niż *poziom p* dla podziału, to dalsze podziały nie są wykonywane i węzeł jest liściem drzewa. Proces jest powtarzany aż do momentu, gdy nie będą możliwe dalsze podziały (przy przyjętych wartościach p dla łączenia i dzielenia). Poprawka Bonferroniego jest stosowana w celu „utrudnienia” uznania za statystycznie istotny wyniku pojedynczego testu, przy wielokrotnym przeprowadzaniu testów na podstawie tych samych danych.

2. Badania preferencji komunikacyjnych

Prezentowane badanie ankietowe, jakie zostało przeprowadzone w grudniu 2010 roku oraz w styczniu 2011 roku. Arkusz ankiety został sformułowany w ten sposób, by pozwolił na określenie preferencji w zakresie sposobów pokonywania przestrzeni miejskiej w z góry zaproponowanych konwencjach: komunikacji zbiorowej, podróży rowerowej (wraz z łączeniem jej z innymi środkami komunikacji – bike and ride) oraz w systemie car pooling. Badaniem objęto całą populację studentów studiów stacjonarnych Państwowej Wyższej Szkoły Zawodowej im. Angelusa Silesiusa w Wałbrzychu (ilość prawidłowo

* Poprawka Bonferroniego jest zwykle dokonywana przez podzielenie poziomu istotności alfa (przyjmowanego zazwyczaj na poziomie 0,05, 0,01, itd.) przez liczbę przeprowadzanych testów.

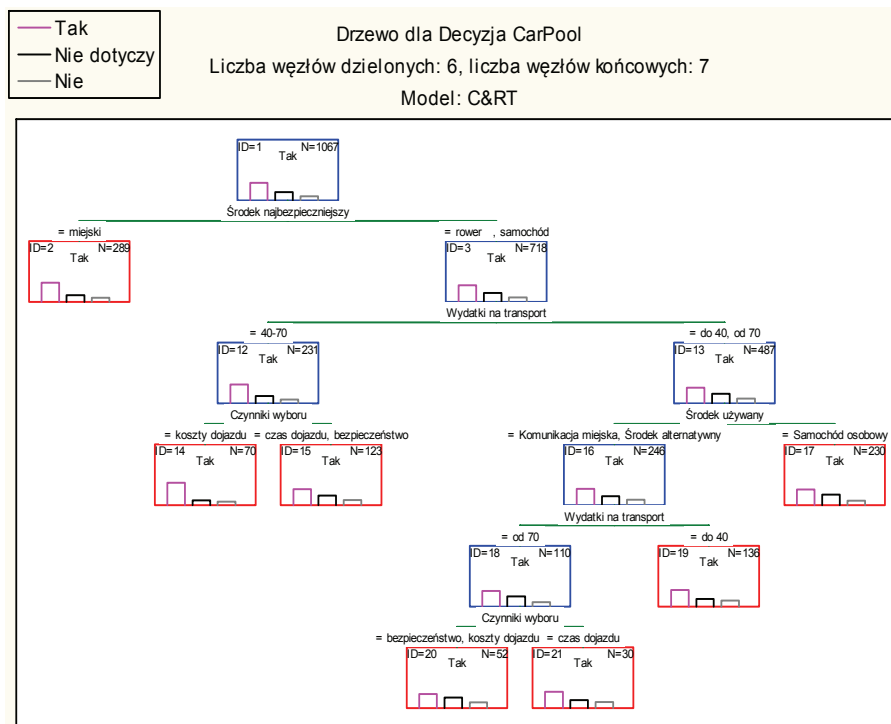
wypełnionych ankiet wyniosła 1070 – zwrot ok. 97%). Badaniu poddano sposób podróżowania na trasie: miejsce zamieszkania – uczelnia – miejsce zamieszkania (grupa podróży obligatoryjnych o wysokim stopniu częstotliwości) i okoliczności (czynniki) mogące wpłynąć na zmianę tego sposobu na inny – alternatywny, bardziej pożądaną z punktu widzenia wymogów (interesów) miasta. Opracowanie wyników struktury udzielonych odpowiedzi [5] pozwoliło na wskazanie podstawowych czynników wpływających znacząco na zmianę dotychczas wykorzystywanego środka transportu (samochodu osobowego) na komunikację miejską*. Czynnikiem mającym znaczenie przy decyzji o podróżowaniu komunikacją zbiorową okazały się zachęty: niskie koszty biletów komunikacji miejskiej oraz większa dostępność komunikacji miejskiej i jeden demotywar dotychczasowego sposobu podróżowania: drastyczny wzrost cen paliwa. Okazało się również, że czynniki, takie jak wysokie opłaty parkingowe czy brak miejsc parkingowych oraz zakaz wjazdu w okolice centrum (gdzie ulokowana jest Uczelnia), nie są znaczącą przeszkodą w korzystaniu z samochodu. Pogłębiona analiza tej konstatacji wykazała, że młodzież nie zwraca uwagi na formalne zakazy uważając, że „nie mają one sensu”, zaś „prawdopodobieństwo wyegzekwowania kary jest znikome”.

Celem kolejnych badań podjętych przez autorów było ustalenie motywatorów i demotywarów w zakresie możliwości udostępniania miejsc w samochodach (Car Pooling) w zależności od zmiennych ograniczających wydatków miesięcznych na paliwo oraz innych czynników wyboru (koszt dojazdu, czas dojazdu, bezpieczeństwo). Kolejno zbadano motywacje przy decyzji o wyborze metody Bike&Ride jako sposobu podróżowania, czyli możliwości podróżowania rowerem, który następnie można zabrać ze sobą do środków komunikacji miejskiej. Decyzję uzależniono od: długości pokonywanej drogi (w km) oraz zainteresowania alternatywnymi środkami komunikacji. Ostatnim elementem badań była analiza motywacji przy decyzji o wyborze roweru jako sposobu podróżowania. Decyzję uzależniono od: długości pokonywanej drogi (w km) oraz innych czynników wyboru (koszt dojazdu, czas dojazdu, bezpieczeństwo). Okazało się, że przyjęte do analizy zmienne nie motywują do podjęcia decyzji o podróżowaniu rowerem. Decyzję o wyborze alternatywnej metody podróżowania odniesiono następnie do możliwości korzystania z komunikacji miejskiej [6].

* Nazywaną tu zamiennie komunikacją zbiorową lub publiczną, mając świadomość pewnych różnic znaczeniowych tych pojęć.

2.1. Analiza preferencji komunikacyjnych z wykorzystaniem modeli drzew C&RT

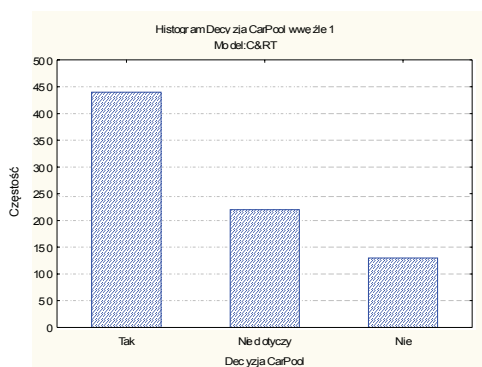
W tej części analizy prezentujemy wyniki modeli drzew C&RT. Drzewa klasyfikacyjne zbudowano kolejno dla różnych alternatywnych metod podróżowania traktując te metody jako zmienne zależne. Decyzję o zmianie środka transportu od uzależniono od następujących czynników: od ponoszonych wydatków na transport, od wykorzystywanych środków transportu, od oceny środka transportu jako bezpieczny oraz od pozostałych czynników wyboru – waloryzacji usług transportowych (koszt dojazdu, czas dojazdu, bezpieczeństwo).



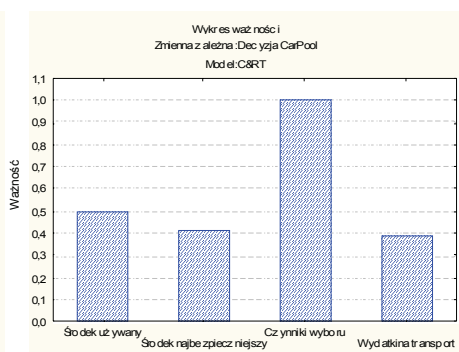
Rys. 2. Drzewo C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania Car Pooling

Przyjmując podjęcie decyzji o wyborze metody Car Pooling jako sposobu podróżowania jako zmienną zależną zbudowano drzewo typu C&RT (rys. 2). Bezpieczny sposób podróżowania postrzegany jako tożsamy z komunikacją miejską nie jest powszechną opinią.

Większość badanych osób (718 osób) wskazała na inne sposoby podróżowania (samochód, rower). W węźle drugim zmienną klasyfikującą był poziom wydatków na transport. W grupie osób o środkowym przedziale wydatków (231 osób) przeważająca grupa ceni sobie przede wszystkim bezpieczeństwo (123 osoby). W pozostałej grupie wyodrębniła się wyraźnie grupa podróżująca jedynie samochodem (230 osób). Szczegółowy histogram rozkładu odpowiedzi w węźle pierwszym (przewaga odpowiedzi „tak”) przedstawia rys. 3. Wykres ważności czynników wśród badanych osób świadczy o tym, że najistotniejszym były czynniki wyboru związany z waloryzacją usługi przewozowej (rys. 4).

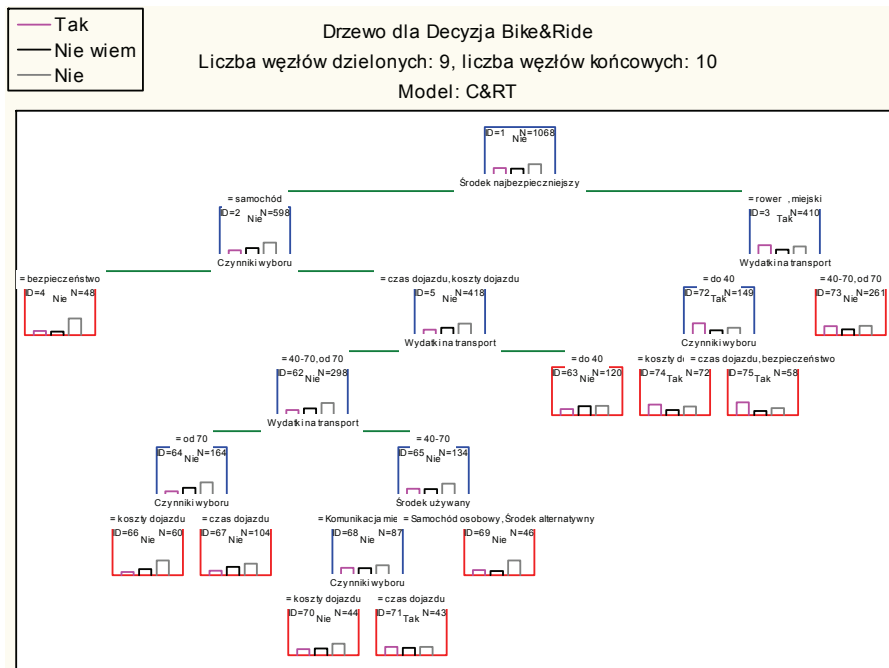


Rys. 3. Histogram w węźle 1 modelu C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania Car Pooling

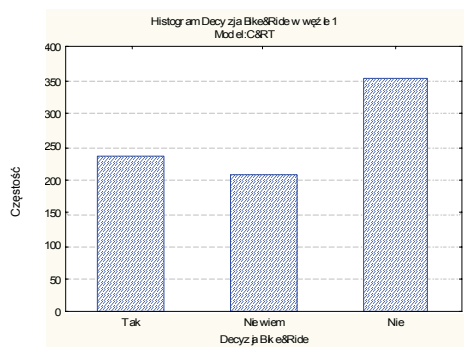


Rys. 4. Wykres ważności wpływu czynników modelu C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania Car Pooling

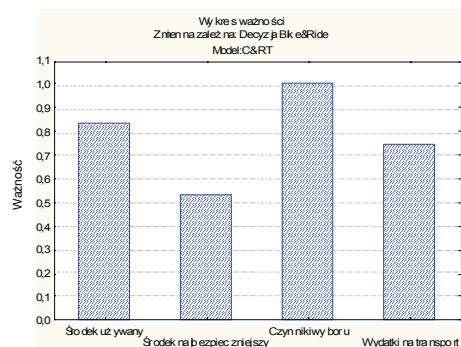
Drzewo typu C&RT zbudowano następnie przyjmując podjęcie decyzji o wyborze metody Bike&Ride jako sposobu podróżowania jako zmienną zależną (rys. 5). Klasyfikacja udzielonych odpowiedzi jest znacząco odmienna od modelu poprzedniego. Szczegółowy histogram rozkładu odpowiedzi w węźle pierwszym (przewaga odpowiedzi „nie”) przedstawia rysunek 6. Nie ma wyraźnego zainteresowania tym sposobem podróżowania. Bezpieczny sposób podróżowania postrzegany jest jako podróż samochodem (598 osób). Wśród pozostałych badanych osób decyzję o zmianie sposobu podróżowania podejmą osoby o najniższych wydatkach na transport. Wykres ważności czynników wśród badanych osób świadczy o tym, że najistotniejszym były czynniki wyboru a następnie aktualnie wykorzystywany środek transportu (rys. 7).



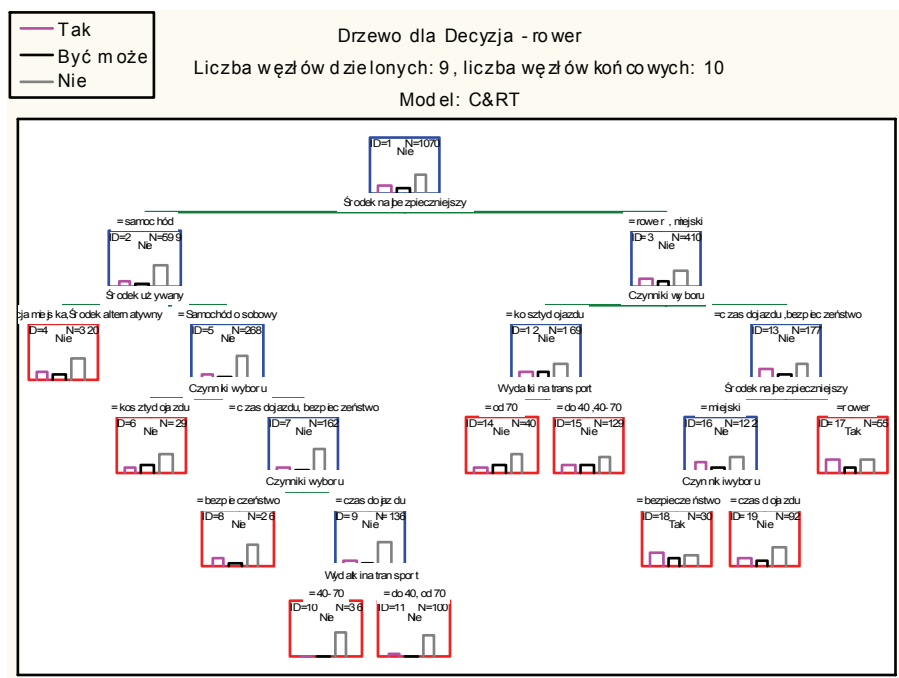
Rys. 5. Drzewo C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania Bike&Ride



Rys. 6. Histogram w węźle 1 modelu C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania Bike&Ride

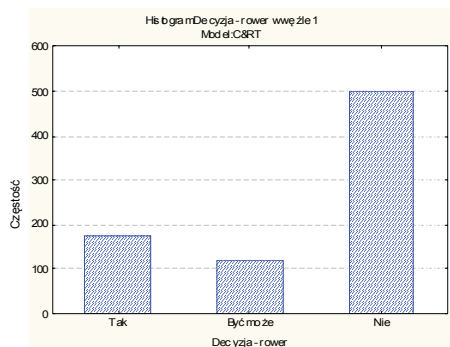


Rys. 7. Wykres ważności wpływu czynników modelu C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania Bike&Ride

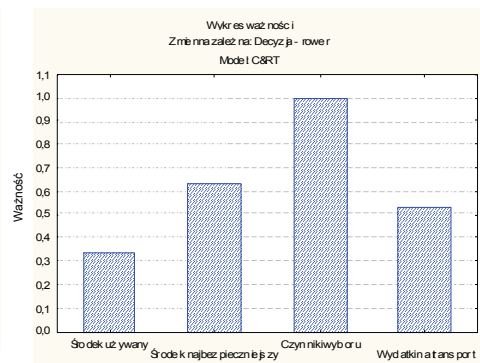


Rys. 8. Drzewo C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania rower

Decyzja o wyborze roweru jako sposobu podróżowania była kolejną rozważoną zmienną zależną (rys. 8). Klasyfikacja udzielonych odpowiedzi jest podobna modelowi poprzedniego. Szczegóły histogram rozkładu odpowiedzi w węźle pierwszym (przewaga odpowiedzi „nie”) przedstawia rys. 9. Nie ma zainteresowania tym sposobem podróżowania. Bezpieczny sposób podróżowania zdaniem respondentów to samochód (595 osób). Wśród badanych czterech czynników największe znaczenie miały pozostałe czynniki wyboru (rys. 10).



Rys. 9. Histogram w węźle 1 modelu C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania rower

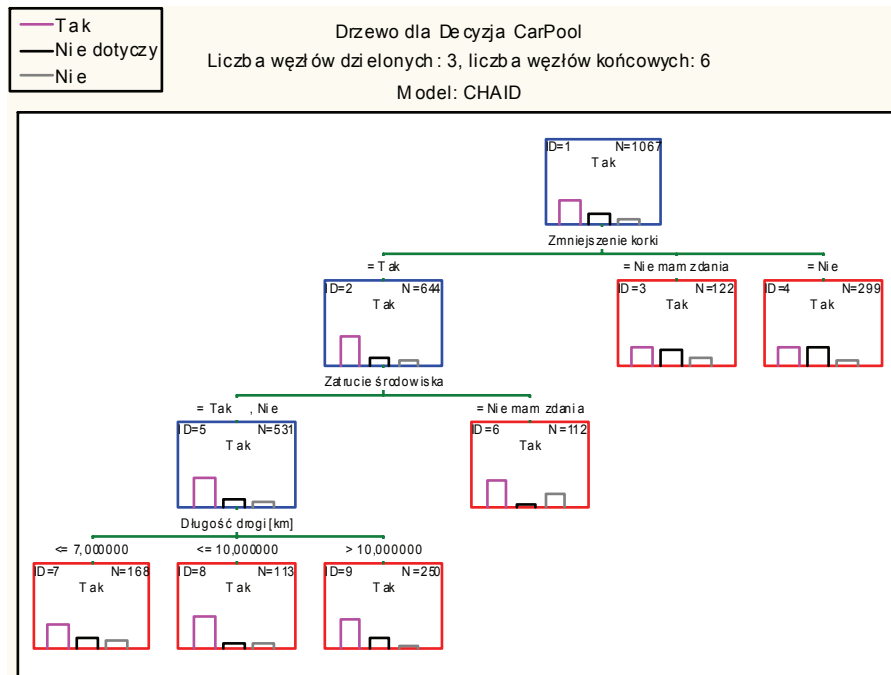


Rys. 10. Wykres ważności wpływu czynników modelu C&RT dla zmiennej zależnej: decyzja o wyborze metody podróżowania rower

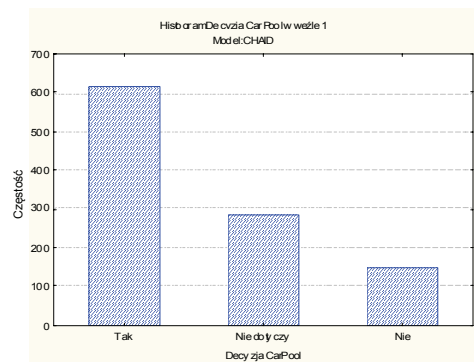
2.2. Analiza preferencji komunikacyjnych z wykorzystaniem modeli drzew CHAID

W tej części analizy prezentujemy wyniki modeli drzew CHAID. Drzewa klasyfikacyjne zbudowano kolejno dla różnych alternatywnych metod podróżowania traktując te metody jako zmienne zależne. Decyzję o zmianie środka transportu od uzależniono od zmiennej ilościowej długość drogi oraz od zmiennych jakościowych zatrucie środowiska, zmniejszenie korków oraz znajomości innych koncepcji podróżowania.

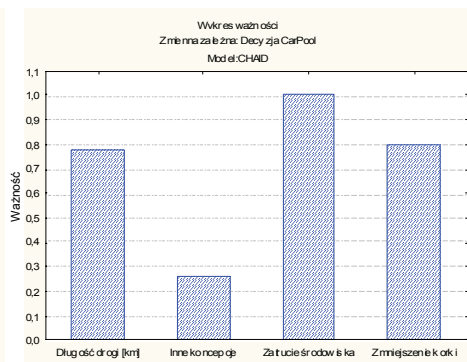
Analizując wyniki dla zmiennej decyzja o wyborze metody Car Pooling jako sposobu podróżowania (zmienna zależna) stwierdzamy (rys. 11), że dla większości respondentów (644 osoby) czynnikiem motywującym jest zmniejszenie natężenia ruchu, a następnie ograniczenie zatrucia środowiska. Szczegółowy histogram rozkładu odpowiedzi w węźle pierwszym (przewaga odpowiedzi „tak”) przedstawia rys. 12. Wśród badanych czterech czynników największe znaczenie miały zatrucie środowiska oraz długość drogi (rys. 13).



Rys. 11. Drzewo CHAID dla zmiennej zależnej: decyzja o wyborze Car Pooling jako metody podróżowania

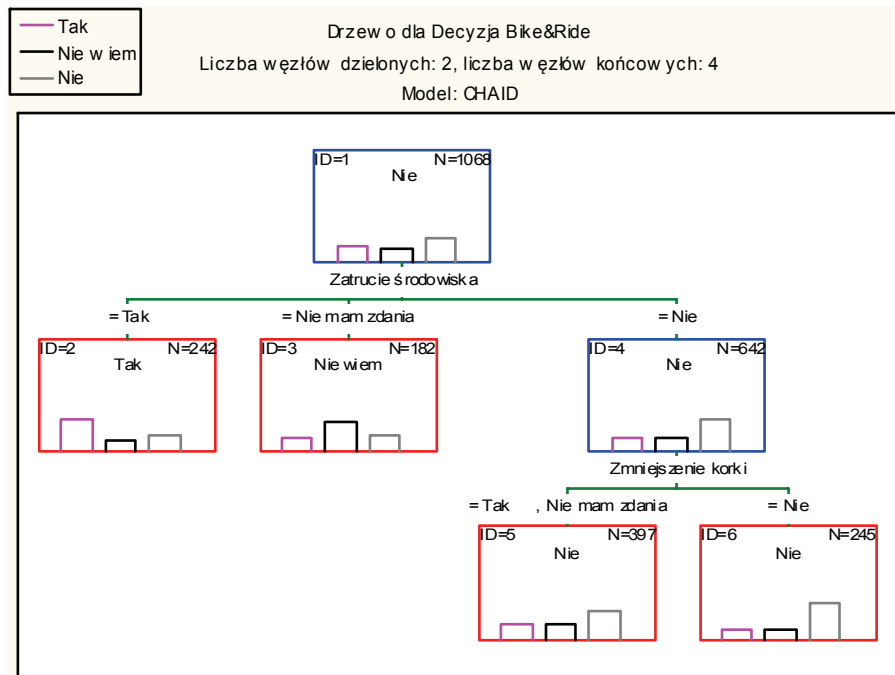


Rys. 12. Histogram w węźle 1 modelu CHAID dla zmiennej zależnej: decyzja o wyborze metody podróżowania Car Pooling

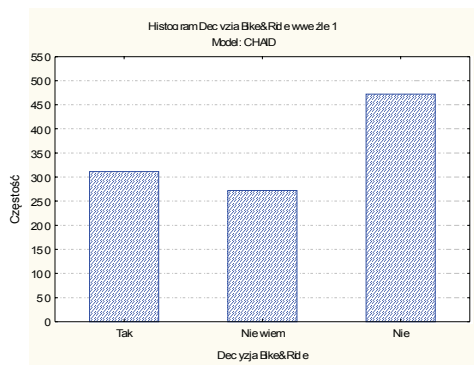


Rys. 13. Wykres ważności wpływu czynników modelu CHAID dla zmiennej zależnej: decyzja o wyborze metody podróżowania Car Pooling

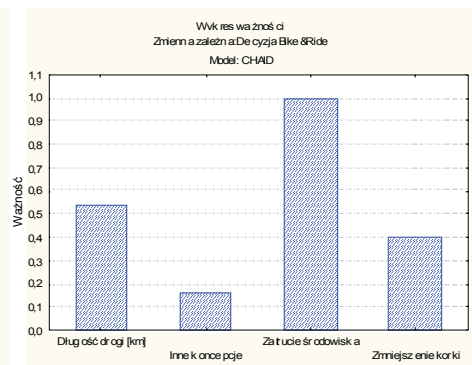
Analizując wyniki dla zmiennej decyzja o wyborze metody Bike&Ride (rys. 14) jako sposobu podróżowania (zmienna zależna) stwierdzamy, że wśród zwolenników tej metody czynnikiem motywującym jest jedynie ograniczenie zatrucia środowiska (242 osoby). Szczegółowy histogram rozkładu odpowiedzi w węźle pierwszym (przewaga odpowiedzi „nie”) przedstawia rys. 15. Ważność czynników jak w poprzednim modelu. Zmienna zależna nie determinuje rankingu ocen czynników. Wśród badanych czterech czynników największe znaczenie miały zatrucie środowiska oraz długość drogi (rys. 16).



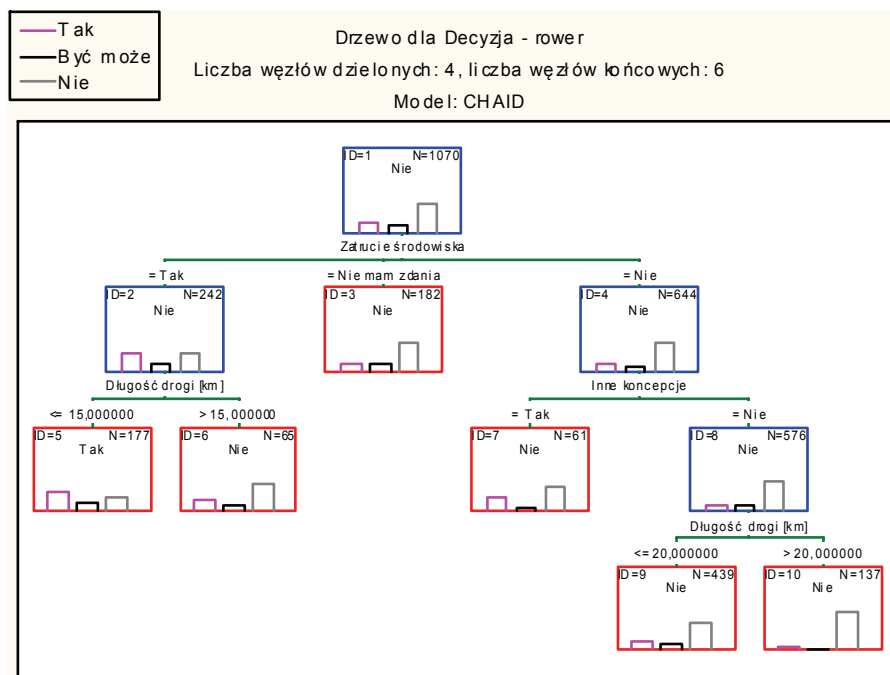
Rys. 14. Drzewo CHAID dla zmiennej zależnej: decyzja o wyborze Bike&Ride jako metody podróżowania



Rys. 15. Histogram w węzle 1 modelu CHAID dla zmiennej zależnej: decyzja o wyborze metody podróżowania Bike&Ride

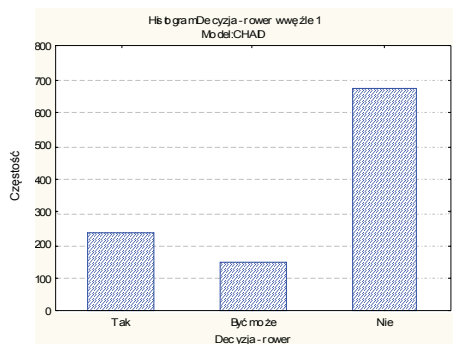


Rys. 16. Wykres ważności wpływu czynników modelu CHAID dla zmiennej zależnej: decyzja o wyborze metody podróżowania Bike&Ride

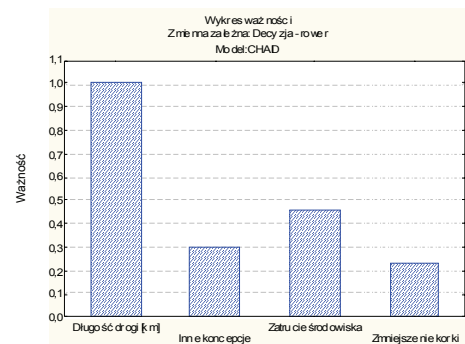


Rys. 17. Drzewo CHAID dla zmiennej zależnej: decyzja o wyborze roweru jako metody podróżowania

Ostatni model przedstawia wyniki dla zmiennej decyzja o wyborze metody podróż rowerem (rys. 17) jako sposobu podróżowania (zmienna zależna). Zmianę sposobu podróżowania podejmą respondenci o odległości nie większej od uczelni niż 15 km (177 osoby). Jest o jedyny czynnik motywujący, ale też wyraźne ograniczenie dla zmiany sposobu podróżowania. Szczegółowy histogram rozkładu odpowiedzi w węźle pierwszym (przewaga odpowiedzi „nie”) przedstawia rys. 18. Ważność czynników odmienna od pozostałych zmiennych zależnych, czyli dwóch wcześniejszych modeli. Zmienna zależna nie determinuje rankingu ocen czynników. Wśród badanych czterech czynników największe znaczenie miał długość drogi (rys. 19).



Rys. 18. Histogram w węźle 1 modelu CHIAD dla zmiennej zależnej: decyzja o wyborze metody podróżowania rower



Rys. 19. Wykres ważności wpływu czynników modelu CHIAD dla zmiennej zależnej: decyzja o wyborze metody podróżowania rower

Podsumowanie

W odróżnieniu od fizyki, gdzie teorie nie mogą zmieniać zjawisk, których dotyczą, w ekonomii jest to możliwe. Co więcej, w naukach społecznych teoria nie musi być prawdziwa, aby wpływać na ludzkie zachowania [4]. Dzieje się tak, gdyż ludzie zdają się mieć kłopoty w odróżnieniu praw opartych na identyczności, od praw statystycznych lub od zupełnego szumu. Dlatego modelowanie zachowań komunikacyjnych, czy raczej ich kształtowanie, nie jest kwestią odpowiednio dobranej metody, gwarantującej zawsze zaplanowane z góry wyniki pod warunkiem dokładnego wykonania wszelkich procedur przepisanych w modelu. Kształtowanie zachowań jest oddziaływaniem na ludzi, zatem działaniem, którego rezultaty trudne są do przewidzenia, zaś dobór metod zależny jest od wielu czynników o charakterze dyskretnym losowym. Niemniej jednak u podstaw kształtowania zachowań leży rozpoznanie motywatorów oraz

demotyatorów zmian w pożądanym zakresie kształtowanych zachowań komunikacyjnych, jak też związanych z nimi poglądów czy przekonań. Wyjawieniu tych elementów służyły zaprezentowane w opracowaniu metody, które do potrzeb budowania programów kształtowania zachowań wydają się dobrze dopasowane, a wyniki dają zadowalające rezultaty.

Literatura

1. Borg J. (2011). Perswazja. Sztuka pozytywnego wpływu na ludzi. PWE, Warszawa.
2. Breiman L., Friedman J., Olsen R., Stone C. (1984). Classification and Regression Trees. Wadsworth, Belmont, CA.
3. Kass G.V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. Applied Statistics, 29, s. 119-127.
4. Soros G. (1999). Kryzys światowego kapitalizmu. Warszawskie Wydawnictwo Literackie Muza S.A., Warszawa.
5. Szołtysek J., Jeziorski P., Trzpiot G. (2011). Analiza uwarunkowań podejmowania decyzji o wyborze alternatywnych sposobów realizacji podróży miejskich. Logistyka. Współczesne wyzwania cz. 2. Wydawnictwo PWSZ im. Angelusa Silesiusa, Wałbrzych.
6. Szołtysek J., Trzpiot G. (2011). Preferencje komunikacyjne studentów jako przesłanki kształtowania programów mobilnościowych. Transport Miejski i Regionalny, 4, s. 28-33.
7. Trzpiot G., Ganczarek-Gamrot A. (2009). Decision Trees for a Virtual Supply Chain. Ekonometria 25, Uniwersytet Ekonomiczny, Wrocław, s. 92-104.

DECISION TREES IN ANALYSIS OF COMMUNICATIONS PREFERENCES

Summary

Most of existing methods of travelling in modern cities are ineffective and not economic, especially when we look on them from congestion creating point of view. In this paper we try to verify a set of determinants, which influence can be the most important in making decision process on changing the method of travelling in the city. As a research method we choose classification trees: C&RT tree and CHAID tree. We consider some of alternative travel method: *Car Pooling*, bike, and *Bike & Ride* versus private cars only. Change of some behaviour in travelling preferences depend on knowledge of alternative travel methods in the city areas. In this paper we use classification trees for making analysis factors which determining travelling preferences, especially decision on changing method of travel for alternative one. The established preferences will be the base of educational and promotional programmes which aim to change the way of citizen travelling methods.

Tomasz Wachowicz

Uniwersytet Ekonomiczny w Katowicach

ANALIZA I WSPOMAGANIE NEGOCJACJI DYSTRYBUCYJNYCH Z ELEMENTAMI ZACHOWAŃ INTEGRACYJNYCH*

Wprowadzenie

Od początku lat osiemdziesiątych ubiegłego stulecia, odkąd to ukonstytuowała się odrębna dyscyplina naukowa zwana analizą negocjacyjną [10], negocjacje poddawane są intensywnym badaniom naukowym zmierzającym w kierunku uchwycenia i opisanie mechanizmów procesu negocjacyjnego, sposobów zachowania negocjatorów oraz możliwości wpływu na osiągnięte przez nich wyniki i optymalizowania tegoż wyniku [6; 8]. W ogólności prace badające proces negocjacyjny i specyfikę zachowań uczestników negocjacji analizują odrębnie dwa specyficzne typy negocjacji związane z całkowicie odmiennymi problemami negocjacyjnymi, których dotyczą. Mowa tu o negocjacjach integracyjnych i dystrybucyjnych [17]. Wyczerpującą analizę procesów negocjacji wpisujących się w powyższą, prostą taksonomię znaleźć można w pracach [2; 8; 12; 13]. Negocjacje integracyjne, będące najczęściej opisywanym w literaturze typem problemów negocjacyjnych, to rodzaj negocjacji integrujących wiele kwestii negocjacyjnych powalających na kompensację poziomów ich rozwiązań (analizę trade-offs), a przez to uzyskanie dwustronnie (wielostronnie) satysfakcjonującego wyniku. Negocjatorzy postrzegani są tutaj jako partnerzy, zmierzający do wypracowania wspólnie akceptowalnego rozwiązania [10]. Posługując się najczęściej wykorzystywanym w teorii negocjacji przykładem, można powiedzieć, iż strony negocjacji integracyjnych koncentrują się na odpowiednim powiększeniu tortu negocjacyjnego (odpowiedniej eksploracji

* Praca naukowa finansowana ze środków na naukę w latach 2009-2012 jako projekt badawczy Ministerstwa Nauki i Szkolnictwa Wyższego numer N N111 234936.

problemu decyzyjnego/negocjacyjnego) tak, aby starczyło go dla każdego z nich. Inaczej problem negocjacji wygląda w sytuacji dystrybucyjnej. Negocjacje dystrybucyjne (pozycyjne) to typ negocjacji dotyczących zazwyczaj podziału pojedynczego dobra. Strony prowadzą zatem dyskusje wokół jednej kwestii (w negocjacjach handlowych zazwyczaj ceny), a ich zadaniem jest ustalenie akceptowalnego przez nie poziomu rozwiązania tej kwestii. Strona, która godzi się na pewne ustępstwa ma świadomość, iż nie może ich skompensować odpowiednimi żądaniami na innych kwestiach negocjacyjnych (jak to ma miejsce w przypadku negocjacji integracyjnych). Zyski jednego negocjatora utożsamiane są więc ze stratami jego partnera (gra o sumie stałej). Mówiąc obrazowo, w przeciwieństwie do negocjacji integracyjnych, w przypadku negocjacji dystrybucyjnych negocjatorzy koncentrują się po prostu i tylko na zapewnieniu sobie jak największego kawałka tortu negocjacyjnego.

Takie podejście do negocjacji dystrybucyjnych sprawia, iż nie jest łatwo poddawać je analizie, a w obliczu czystej konkurencji naznaczającej stosunki między stronami, trudno również proponować pewne narzędzia wspomagania takiego procesu negocjacyjnego. Nie sprzyja temu antagonistyczne nastawienie partnerów wobec siebie, którzy niechętnie (jeśli nie wcale) przekazują stronom trzecim (potencjalnie wspomagającym cały proces negocjowania) prawdziwe informacje o jedynym strategicznym znaczeniu w tych negocjacjach, tj. poziomach rezerwacji i aspiracji, na podstawie których można by dokonać analizy i predykcji „sprawiedliwego” kompromisu. Po drugie zaś, zamkniętość na pojedynczą kwestię negocjacyjną nie pozwala na rozwój pewnych mechanizmów proaktywnych w zakresie poszukiwania alternatywnych propozycji kompromisu wobec impasu, jaki zaistnieć może w toku rozmów. Przeszukując przestrzeń negocjacyjną nie można bowiem prowadzić analizy trade-off i kompensować stronom słabych wyników pozostałymi kwestiami, gdyż tych kwestii nie ma*. Można byłoby jednak stosować narzędzia, o których mowa wyżej, gdyby w toku negocjacji dystrybucyjnych pojawiły się jednostronnie (lub z inicjatywy obydwu stron konfliktu) pewne działania integracyjne, zmierzające do wprowadzenia do negocjacji dodatkowych elementów podlegających dyskusjom, tj. rozszerzające wyjściowy problem negocjacyjny, jakkolwiek nie zmieniające wyjściowej kwestii negocjacyjnej. Takie spojrzenie na negocjacje dystrybu-

* Analiza taka prowadzona w negocjacjach integracyjnych niejednokrotnie pozwala na zidentyfikowanie wariantu dominującego aktualnie dyskutowane rozwiązanie, czyli poprawę rozwiązania z punktu widzenia wszystkich zainteresowanych stron jednocześnie.

cyjne było już przytaczane w literaturze [6; 13]. Jednak pytanie o to, czy takie działania integracyjne rzeczywiście pojawiają się w toku negocjacji dystrybucyjnych pozostaje kwestią otwartą, wymagającą głębszych badań empirycznych.

Niniejsza praca przybliży problematykę negocjacji dystrybucyjnych, a jednocześnie stara się pokazać pewne alternatywne metody ich analizy. W pierwszej części pracy, mającej charakter jedynie systematyzujący dotychczasową wiedzę z tego zakresu, omówiono podstawowe zagadnienia związane z analizą negocjacji dystrybucyjnych: cele działań facylitacyjnych, modelowanie formalne, mechanizmy i procedury wspomagania i wynikające z nich problemy. W drugiej części pracy przedstawiona zostanie autorska idea podejścia do negocjacji dystrybucyjnych, w których zakłada się istnienie wątków integracyjnych między negocjatorami. Punktem wyjścia w formułowaniu takiej idei były przeprowadzone przez autora pracy eksperymenty negocjacyjne (negocjacji dwustronnych), z których wynikało, iż w znacznym odsetku negocjacji dystrybucyjnych pojawiają się zachowania integracyjne jeśli nie obydwu, to przynajmniej jednej ze stron negocjacji. Analizując działanie negocjatorów zaproponowano procedurę analizy takich negocjacji, bazującą na metodzie even swaps [5]. Procedura bazuje na zmodyfikowanym uprzednio algorytmie even swaps wykorzystywanym do wspomagania dwustronnych negocjacji integracyjnych [14]. Wykorzystanie go w negocjacjach dystrybucyjnych wymaga kolejnej modyfikacji i niejako odwrócenia idei wyjściowej kumulacji ocen częściowych w skalarnej ocenie globalnej na rzecz dekumulacyjnego (subtrakcyjnego) podejścia do tworzenia systemu oceny ofert. Procedura ta zostanie omówiona w ostatniej części pracy.

1. Negocjacje dystrybucyjne i ich wspomaganie

Negocjacje dystrybucyjne opisywane są w literaturze jako negocjacje typu win-lose [4] negocjacje o sumie zero [10] negocjacje nad ustalonym tortem (fixed-pie) [13] czysto konfliktowe lub konkurencyjne. Przygotowując się do negocjacji dystrybucyjnych, w których to rozmowy stron dotyczą zazwyczaj tylko pojedynczej kwestii negocjacyjnej (lub ograniczonych zasobów) i zmierzają do ustalenia wspólnie akceptowalnego poziomu rozwiązania tej kwestii, negocjatorzy koncentrują się na zgoła odmiennych zagadnieniach i problemach, niż ma to miejsce w przypadku negocjacji integracyjnych. Nie jest tu aż tak

istotna analiza decyzyjna potencjalnych kompromisów i ich konsekwencji w świetle przyjętych przez strony preferencji. Nie jest istotna analiza jakości kompromisów, a w ogóle niemożliwa staje się typowa w przypadku negocjacji integracyjnych analiza kompensacji (trade-off analysis). Aby zrozumieć istotę negocjacji dystrybucyjnych przedstawimy poniżej syntetyczny, uproszczony opis takiej sytuacji*. Przyjmujemy, iż analizujemy proces negocjacji dwustronnych, w których dwoje negocjatorów (sprzedawca i nabywca) ma ustalić wspólnie akceptowalną wartość deterministyczną pewnej zmiennej ciągłej (wyrażonej np. w jednostkach pieniężnych).

Przed przystąpieniem do negocjacji negocjatorzy powinni zdefiniować swoje BATNA** [4]. Jest to powszechnie przyjęty w praktyce negocjacyjnej element analizy pre-negocjacyjnej (przygotowawczej) polegający na rozpoznaniu możliwości rozwiązania problemu negocjacyjnego poza samymi negocjacjami z partnerem, np. poprzez znalezienie ekwiwalentu przedmiotu negocjacji dostępnego u innych dostawców na rynku. Identyfikacja takiego ekwiwalentu pozwala negocjatorom na zdefiniowanie poziomów zastrzeżonych (cen rezerwacji) – reservation values – wyznaczających minimalne progi akceptacji negocjowanej kwestii, oznaczone kolejno przez s dla sprzedawcy i b dla nabywcy. Innymi słowy zakłada się, iż żaden z negocjatorów nie zaakceptuje kompromisu (x^*) na poziomie gorszym od własnego poziomu zastrzeżonego $s(b)$. Im większy jest x^* od s , tym większa tzw. nadwyżka sprzedawcy (seller's surplus) – $x^* - s$. Konsekwentnie dla poziomu zastrzeżonego b kupującego, jeśli $b \geq x^*$, wówczas jego nadwyżka (buyer's surplus) wyniesie $b - x^*$.

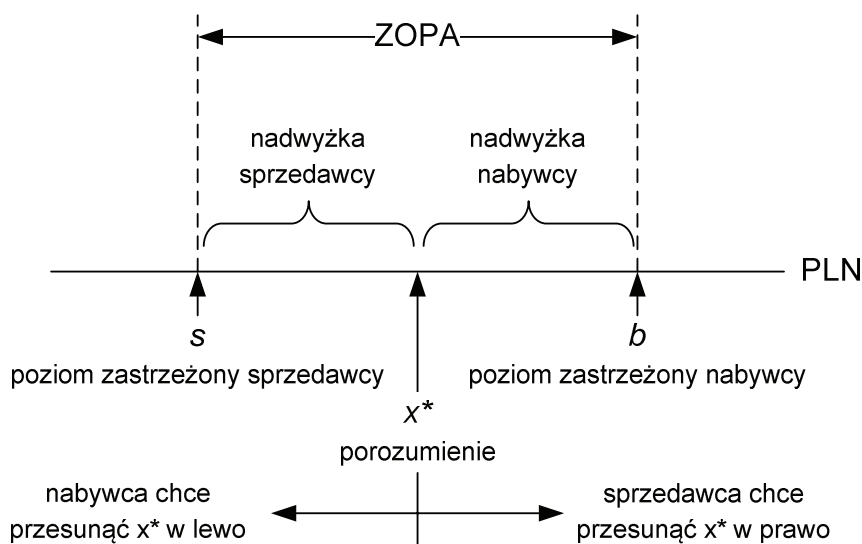
Można się zatem łatwo zorientować, iż aby istniała szansa na kompromis w dwustronnych negocjacjach dystrybucyjnych poziom zastrzeżony kupującego, tj. maksymalna cena, jaką jest w stanie zapłacić za negocjowany przedmiot, musi być większy (lub równy) poziomowi zastrzeżonemu sprzedającego, tj. minimalnej cenie, za którą chce on sprzedać ów przedmiot. Formalnie, musi zachodzić $b > s$. Jeśli tak jest mówimy, że istnieje niepusta strefa potencjalnego kompromisu ZOPA (Zone of Possible Agreement)***. ZOPA jest zatem przedziałem liczbowym $\langle s; b \rangle$, a jej wielkość mierzy się różnicą $b - s$. Jeśli więc

* W opisie sytuacyjnym wzorowano się na pracy [10].

** BATNA jest akronimem angielskiego sformułowania Best Alternative To Negotiated Agreement.

*** Strefa potencjalnego kompromisu (ZOPA) w polskojęzycznej terminologii nazywana jest również obszarem targowania się [12].

negocjatorzy zgodzą się na pewien kompromis x^* leżący w strefie kompromisu ZOPA, wówczas uzyskują wspomniane wcześniej nadwyżki $x^* - s$ oraz $b - x^*$. Suma tych nadwyżek jest równa wielkości ZOPA $b - s$ i jest niezależna od wartości samego porozumienia x^* . Opisaną powyżej sytuację negocyjacyjną da się przedstawić graficznie (rys. 1).



Rys. 1. Geometria negocjacji dystrybucyjnych

Źródło: [10, s. 111].

Według przedstawionego wyżej opisu negocjacje dystrybucyjne można by zatem potraktować jako grę o sumie stałej. Nie jest tak jednak do końca, bowiem specyfika negocjacji sprawić może, iż mimo że $b > s$ (ZOPA formalnie istnieje) strony mogą nie dojść do porozumienia, tj. – w skutek wzajemnych nieporozumień lub niechęci do czynienia ustępstw – nie zgodzić się na żaden wynik x^* należący do ZOPA. Dlatego też bardziej uzasadnione jest traktowanie negocjacji jako gry o quasi stałej sumie [10]. Taka perspektywa jest też o tyle uzasadniona, że gracze (negocjatorzy) zazwyczaj też nie znają wielkości całego tortu negocyjacyjnego, tj. wielkości ZOPA $b - s$.

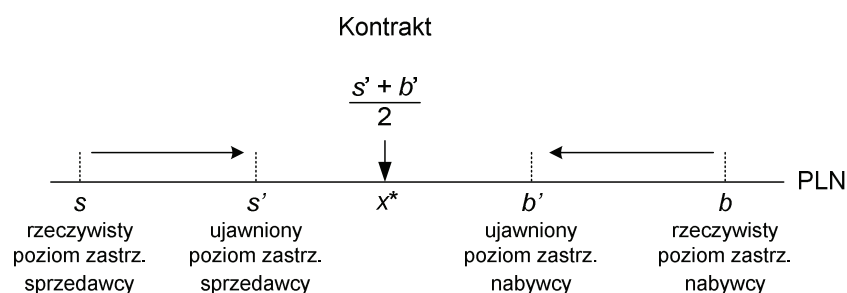
Jak zatem łatwo wywnioskować z powyższego syntetycznego opisu negocjacji dystrybucyjnych, podstawowym zadaniem, przed jakim stają negocjatorzy w negocjacjach dystrybucyjnych, jest próba odpowiedzi na strategiczne z ich punktu widzenia pytania, tj.:

1. Kto powinien rozpocząć negocjacje: otwierając negocjacje przekazujemy partnerowi strategiczną informację o potencjalnych poziomach aspiracji i rezerwacji. Partner, rozpoznając wcześniej w fazie prenegocyjacyjnej strefę potencjalnego kompromisu może na podstawie oferty otwarcia zweryfikować swoją opinię o naszym zakresie żądań i pozycji.
2. Czy i w jaki sposób da się określić poziom zastrzeżony (cenę rezerwacji) partnera: wiedza o poziomie zastrzeżonym partnera pozwala na rozsądne ustalenie własnej oferty otwarcia negocjatora oraz przygotowanie takiej strategii ustępstw, aby doprowadzić do wyniku spełniającego poziom aspiracji negocjatora.
3. Jaka powinna być oferta otwarcia (jak bardzo skrajna w żądaniach): negocjator jest w stanie w swej ofercie początkowej odpowiednio oddalić się od ceny zastrzeżonej lub oferty otwarcia partnera zostawiając sobie nieco miejsca na potencjalne kompromisy w toku negocjacji.
4. Jaka powinna być rozsądna odpowiedź na ofertę otwarcia partnera: mając na uwadze powszechne przekonanie o potencjalnym kompromisie leżącym zazwyczaj w połowie, między ofertami otwarcia obydwu negocjatorów i znając własne poziomy aspiracji i rezerwacji negocjator powinien umieć odpowiednio zareagować na ofertę otwarcia partnera.
5. Jak powinna wyglądać ścieżka ustępstw (strategia ustępstw) na kolejne rundy negocjacji: negocjator przygotowując strategię ustępstw określa po-niekąd samą strategię negocjowania. Musi zdecydować o ile i na jakim eta-pie negocjacji ustąpić partnerowi, demonstrując tym samym pewną postawę negocjacyjną, aby skłonić go do akceptacji założonego przez siebie po-ziomu rozwiązania spornej kwestii, tj. przeforsować satysfakcjonujący dla siebie wynik.

Istnieje wiele metod opisu i wspomagania negocjacji dystrybucyjnych, pozwalających uzyskać odpowiedź na wymienione powyżej pytania [10]. Najprostsze z nich mogą sięgać do analizy SWOT i próby przeprowadzenia na bazie uzyskanych wyników analizy teoriogrowej w kierunku identyfikacji strategii najlepszych odpowiedzi na przewidywaną ofertę otwarcia partnera [18]. Analiza ofert zmierza wówczas do identyfikacji rozkładu prawdopodobieństwa ocen otwarcia, a następnie, wobec braku przesłanek co do racjonalnego wyboru przez partnera któregośkolwiek z poziomów realizacji analizowanej zmiennej losowej, wykorzystania elementów gier z naturą do podejmowania decyzji na podstawie średnich ważonych [10, s. 137-144].

Aby uniknąć nieporozumień w toku wymiany ofert negocjacyjnych i wyeliminować niekorzystne efekty emocji, mogące mieć wpływ na wywoływanie impasów lub co gorsza na zerwanie negocjacji mimo niepustej ZOPA, jako narzędzie wspomagania rozwiązania dystrybucyjnej sytuacji negocjacyjnej proponuje się aukcję dwustronną z jednoczesnym ujawnianiem ofert. Mechanizm tej aukcji polega na jednoczesnym ujawnieniu przez strony swych własnych

poziomów zastrzeżonych i, jeśli istnieje ZOPA, ustaleniu poziomu kontraktu będącego wartością średnią zgłoszonych przez uczestników aukcji ofert. Aby zapewnić ujawnienie rzeczywistych poziomów zastrzeżonych, a nie fałszywych deklaracji, na które wpłynęła otrzymana wcześniej oferta otwarcia partnera, przyjmuje się, że deklaracja ofert następuje poprzez ich zapisanie, umieszczenie w kopercie i złożenie stronie trzeciej (organizatorowi aukcji), który następnie publicznie ogłasza oferty. Niestety i przy takiej organizacji aukcji istnieją obustronne pokusy do nieujawniania przez negocjatorów ich rzeczywistych poziomów zastrzeżonych. Jeżeli bowiem przyjąć, iż partner zgłosi nam w ofercie aukcyjnej swój rzeczywisty poziom zastrzeżony, jako sprzedający skorzystamy na zawyżeniu swojego poziomu deklarując pewne s' zamiast s (gdzie $s' > s$). Analogicznie kupujący zyska na zaniżeniu własnego poziomu zastrzeżonego i deklaracji b' w miejsce b (gdzie $b' < b$). Mechanizm działania tak opisaną aukcji dwustronnej z jednoczesnym ujawnianiem ofert zobrazowano graficznie na rys. 2. .



Rys. 2. Procedura wyznaczania rozwiązania w aukcji dwustronnej z jednoczesnym ujawnianiem ofert

Źródło: [10, s. 123].

Pierwsze spostrzeżenie, jakie od razu nasuwa się wobec tak zdefiniowanej procedury rozstrzygania aukcji dwustronnej, dotyczy potencjalnej nieefektywności tego narzędzia. Jeśli bowiem przyjmiemy, że uczestnicy aukcji mają tendencję do przesadnego zawyżania (zaniżania) swoich poziomów zastrzeżonych w deklarowanych ofertach aukcyjnych, wówczas może się okazać, iż mimo istnienia ZOPA aukcja zakończy się fiaskiem na skutek tychże wyolbrzymień, wobec $s' > b'$. Taką tendencję w działaniu negocjatorów, a co za tym idzie niebezpieczeństwo fiaska aukcji wspomagającej negocjacje dystrybucyjne, potwierdza eksperyment Chatterjee i Samuelsona (1981) dotyczący teoriogrowej analizy równowagi strategii nabywcy i sprzedawcy w sytuacji, gdy przyjmuje się założenie o jednostajnym rozkładzie prawdopodobieństwa nad wartością poziomu zastrzeżonego partnera w znanym przedziale zmienności. Eksperyment ten powtórzony 20 lat później przez Howarda Raiffę [10]

potwierdza, że strategie działania negocjatorów w tak zdefiniowanej aukcji z prawdopodobieństwem 0,46 prowadzą do pustej ZOPA. Pojawia się zatem pytanie, czy nie można próbować stosować innych narzędzi wspomagania procesu negocjacji dystrybucyjnych i czy rzeczywiście konieczne jest utrzymanie silnego założenia o całkowitej dystrybucyjności tego procesu, zmuszające do poszukiwania rozwiązań jedynie w obszarze narzędzi teoriogrowych.

2. Dystrybucyjny eksperyment negocjacyjny

2.1. Opis badania

Przedstawione w poprzednim rozdziale opis i metody analizy negocjacji dystrybucyjnych odpowiadają klasycznemu podejściu do modelowania tego typu sytuacji konfliktowych. W istocie, przy założeniu stricte dystrybucyjnej sytuacji konfliktowej omówione podejście wydaje się być uzasadnione, a ponadto w zakresie potencjalnej problematyki związanej z negocjowaniem w warunkach przeciwstawnych interesów wyczerpuje zestaw możliwych analiz. Podejście to może jednak okazać się niewystarczające, jeśli w warunkach negocjacji dystrybucyjnych pojawią się pewne integracyjne zachowania negocjatorów, zmieniające *de facto* wyjściowy problem negocjacyjny. Aby zbadać rzeczywisty sposób zachowania negocjatorów, uczestników negocjacji dystrybucyjnych, autor niniejszej pracy zaprojektował własne badanie symulacyjne, polegające na rozegraniu dwustronnych, dystrybucyjnych negocjacjach handlowych według ustalonego scenariusza.

Głównym celem zaprojektowanego eksperymentu było uchwycenie sposobu rozgrywania negocjacji handlowych przez nabywców i odbiorców w negocjacyjnym problemie dystrybucyjnym. Jednocześnie zamierzano zweryfikować efektywność standardowego narzędzia wspomagania negocjacji dystrybucyjnych, tj. aukcji dwustronnej oraz porównać wyniki osiągnięte za pomocą tego typu aukcji z wynikami ustalonymi w drodze negocjacji. Na cele poboczne badania składały się również weryfikacja hipotezy o przewidywalnym wyniku negocjacji dystrybucyjnych, kształtującym się w połowie [10] między ofertami otwarcia obydwu stron oraz zbadanie umiejętności przewidywania przez negocjatorów poziomów zastrzeżonych swoich partnerów oraz dopasowania do nich własnych ofert otwarcia.

Problem negocjacyjny w przeprowadzonym badaniu dotyczył zakupu/zbycia używanego samochodu osobowego i był rozszerzeniem pewnego znanego badania, przeprowadzonego wcześniej przez Hammonda i Raiffę [10] jedynie w celu identyfikacji ofert otwarcia stron i ich wpływu na końcowy wynik, jaki strony ustalą w procesie negocjowania. Jako że badanie kierowane było do osób młodych, studentów studiów dziennych Uniwersytetu Ekonomicznego w Katowicach, na przedmiot transakcji wybrano niedrogi, mały

samochód – 12-letniego Fiata Seicento. Stosunkowo niska średnia cena giełdowa tego samochodu (będąca w finansowym zasięgu młodych studentów) miała na celu urealnić przedmiot transakcji i bardziej zaangażować uczestników w eksperyment. Dla osób odgrywających rolę nabywcy i sprzedawcy przygotowane zostały karty opisu problemu zawierające publiczne (wspólnie znane) informacje na temat samego problemu oraz strategiczne informacje prywatne (znane osobno stronom), dotyczące danych szczegółowych przedmiotu transakcji, potencjalnej przestrzeni negocjacyjnej oraz sugerowanych poziomów rezerwacji. W obydwu scenariuszach, powołując się na dane obiektywne (notowania giełdowe, ogłoszenia internetowe itp.), identycznie zdefiniowano przestrzeń negocjacyjną, która obejmowała zakres cen od 1700 do 6500 zł. Dodatkowo przekazane stronom informacje dotyczące ich indywidualnych poziomów zastrzeżonych (cen rezerwacji), wynikające z subiektywnie zdefiniowanych dla nich ograniczeń, wynoszące odpowiednio 3000 zł dla sprzedawcy i 5500 zł dla nabywcy, wyznaczały niepuścą strefę potencjalnego kompromisu (ZOPA). Jednocześnie scenariusze obydwu negocjatorów niejednoznacznie sugerowały, iż wymienione ceny mogą jeszcze zostać naruszone pod szczególnymi warunkami. Obustronnie satysfakcjonujący kontrakt był zatem możliwy, a jego wartość zależała jedynie od poziomów aspiracji obydwu negocjatorów oraz ich umiejętności negocjacyjnych i skłonności do ustępstw i kompromisów.

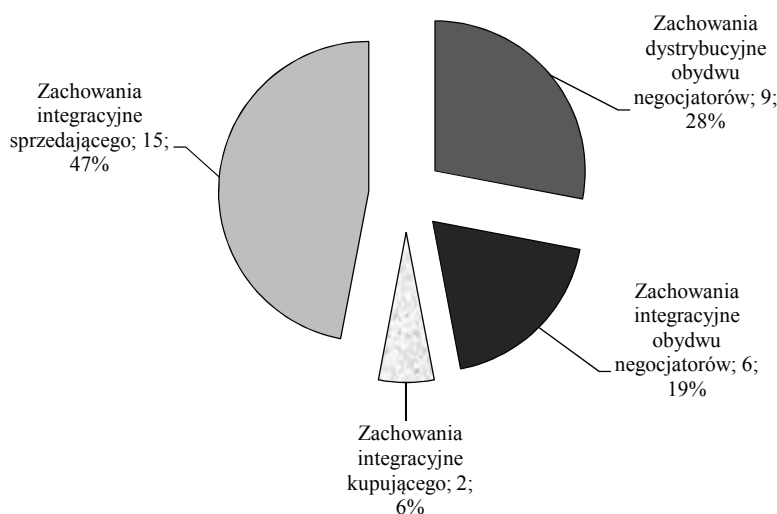
Integralnym elementem badania był indywidualny protokół negocjacyjny każdego uczestnika eksperymentu, opracowany w postaci ankiety badawczej zbierającej informacje o negocjatorach i ich postawie we wszystkich trzech fazach procesu negocjacyjnego. Zakres informacji gromadzonych o fazie pre-negocjacyjnej każdego uczestnika eksperymentu dotyczył postawy (zachowania) w przypadku rozgrywania problemu negocjacyjnego za pomocą narzędzia wspomagającego w postaci aukcji dwustronnej z równoczesnym ujawnianiem ofert zastrzeżonych, estymacji poziomu zastrzeżonego swojego partnera oraz własnego poziomu aspiracji. W ramach danych fazy negocjacji właściwych gromadzone były informacje o argumentacji towarzyszącej kolejnym ofertom negocjacyjnym przedstawianym przez negocjatora. Informacje z fazy pre-negocjacyjnej dotyczyły subiektywnej oceny przez negocjatora atmosfery rozmów, własnej postawy negocjacyjnej oraz postawy partnera. W ramach ostatniego punktu ankiety negocjatorzy proszeni byli o wskazanie ewentualnych oryginalnych rozwiązań proponowanych przez nich indywidualnie w celu przełamania impasu negocjacyjnego i zawiązania satysfakcjonującego kompromisu. Dodatkowo badanie uzupełniał zapis przebiegu negocjacji (transkrypt) specyfikujący kolejność składanych przez negocjatorów ofert i pozwalający śledzić ścieżki ich ustępstw.

W badaniu wzięli udział studenci studiów dziennych Uniwersytetu Ekonomicznego w Katowicach, łącznie 64 studentów kierunków informatyka ekonomiczna (5 roku studiów jednolitych magisterskich), logistyka (1 roku studiów

licencyjnych) oraz studenci zagraniczni z kierunku international business. Studenci utworzyli 32 pary negocjacyjne, które w fazie prenegocjacyjnej zapoznali się z indywidualnym opisem ról oraz wypełnili pierwsze trzy punkty protokołu negocjacyjnego definiując: ofertę negocjacyjną dla aukcji dwustronnej z jednocześnie ujawnianymi ofertami, rozgrywanej równolegle z negocjacjami, przewidywania co do poziomu rezerwacji swojego partnera oraz własny poziom aspiracji. Następnie prowadzili negocjacje zapisując historię ofert (ścieżki ustępstw), by w etapie postnegocjacyjnym wypełnić pozostałe punkty kwestionariusza oraz ustalić wyniki (poziom kontraktu) jakie osiągnęliby rezygnując z negocjacji, a biorąc w zamian udział w aukcji dwustronnej. Autor ma świadomość, że do wyciągania naukowo rzetelnych wniosków badawczych konieczna byłaby znacznie większa grupa badawcza*, niemniej jednak zaobserwowane podczas eksperymentu zjawiska dają podstawę do podjęcia badań nad nowymi, alternatywnymi podejściami do modelowania i wspomagania negocjacji dystrybucyjnych.

2.2. Analiza i interpretacja uzyskanych wyników

Realizując główny cel badania w pierwszym kroku przeprowadzono analizę rzeczywistych zachowań negocjatorów. Strukturę zachowań przedstawia wykres kołowy (rys. 3).



Rys. 3. Struktura zachowań w eksperymencie negocjacji dystrybucyjnych (liczba sesji negocjacyjnych o danym typie zachowań stron)

* Powtórzenie tego badania na większej grupie ankietowanych z bardziej szczegółową ankietą weryfikującą zachowania i motywy tych zachowań stanowiłoby kolejny etap badawczy autora pracy nad negocjacjami dystrybucyjnymi z wątkami integracyjnymi.

Okazuje się, że aż w 23 na 32 rozegranych przypadków negocjacyjnych (72%)* negocjatorzy okazywali zachowania integracyjne proponując rozszerzenie przedmiotu negocjacyjnego przy utrzymaniu oficjalnie ceny, jako jedynego kryterium oceny tegoż przedmiotu. W sześciu przypadkach negocjacyjnych (19%) zachowania integracyjne przejawiały obie strony negocjacji. Zdecydowanie natomiast przeważały zachowania integrujące po stronie sprzedawców pojazdów, w sumie 66% przypadków (licząc 9 przypadków negocjacji gdzie obie strony wykazywały takie zachowania i 15 przypadków, gdzie zachowania integracyjne przejawiali wyłącznie sprzedawcy). Wśród kupujących takie zachowania były zdecydowanie rzadsze, w sumie zaobserwowano je w 25% sesji negocjacyjnych (tj. w 8 przypadkach, a jedynie w 2 z nich zachowania integracyjne przejawiała tylko strona kupująca). Prezentowane przez strony ruchy integracyjne dotyczyły rozszerzenia problemu negocjacji o pewne, nieujęte w wyjściowym opisie sytuacyjnym elementy wyposażenia samochodu lub usługi. Negocjatorzy wykazywali sporą kreatywność w tworzeniu kwestii integracyjnych, czasami proponowali nawet rozszerzenie problemu dystrybucyjnego o więcej niż jedną kwestię. Niemniej najczęściej proponowali komplet dodatkowych opon (7 przypadków), darmowy serwis gwarancyjny** (7 przypadków) oraz drobne upominki: audio – płyty, kasety; kosmetyka samochodowa (8 przypadków). W pozostałych przypadkach propozycje wahały się od skromnego talonu na benzynę po darmowy montaż klimatyzacji. Biorąc pod uwagę rozszerzenie problemu negocjacyjnego o dodatkowe kwestie integracyjne naturalnym wydawałoby się, że wartości kontraktu w przypadku wystąpienia wątków integracyjnych powinny być wyższe, niż w przypadku czystych negocjacji dystrybucyjnych. W istocie w tym pierwszym przypadku zanotowano średnią wartość kontraktu na poziomie 4078,36 zł, podczas gdy w negocjacjach stricte dystrybucyjnych średni kontrakt wyniósł 4017,50 zł. Różnica ta jednak nie jest zbyt duża. Dodatkowo przeprowadzony test ANOVA nie pozwolił na przyjęcie hipotezy o istotnej różności obydwu średnich.

W kolejnym kroku analizy starano się sprawdzić, czy wykorzystanie popularnego narzędzia wspomagającego – dwustronnej aukcji kopertowej – pozwala efektywnie rozwiązywać negocjacje w ramach problemów dystrybucyjnych i jest narzędziem skuteczniejszym (i w jakiej formie) od samego procesu negocjacji. Tymczasem po analizie wyników przeprowadzonych eksperymentów stwierdzono, iż o ile w przypadku negocjacji kompromisem zakończyło się 30 sesji (94%), o tyle procedura aukcyjna doprowadziłaby do kontraktu jedynie w 10 przypadkach (31%). Aż 22 pary negocjacyjne tak przesadnie zawiżyły (zaniżyły) swoje ujawniane wartości poziomów zastrzeżonych

* Oszacowania procentowe wprowadzono jedynie w dodatkowym celu informującym, mając jednocześnie świadomość, iż dla małych próbek (poniżej 100 obserwacji) operowanie wartościami procentowymi może niewłaściwie oddawać rzeczywistą strukturę zjawiska.

** W opisie roli sprzedawcy informowano, iż jest on właścicielem firmy zajmującej się handlem i naprawą samochodów używanych, dysponującym własnym warsztatem samochodowym.

b' i s' , iż skutkowało to pustą strefą ZOPA ($s' > b'$) i brakiem możliwości ustalenia wartości kontraktu. Zaskakująca jest również skala zawyżania (zaniżania) własnych poziomów zastrzeżonych przez uczestników eksperymentu. Zaznacza się przy tym pewna asymetria w postawie dwóch stron. O ile nabywcy zaniżali swój poziom rezerwacji (wynoszący 5500 zł) średnio o 2565,62 zł, o tyle sprzedawcy zawyżali średnio o 600 zł więcej. Ich ujawniany w postaci oferty aukcyjnej średni poziom rezerwacji wyniósł $\bar{s}' = 6155,37$ zł i był wyższy o 3155,37 zł od rzeczywistego (3000 zł). Takie rezultaty wskazywać mogą na zbyt nonszalanckie podejście do problemu i partnera, zapewne wyniesione z doświadczeń negocjacyjnych, w których przesadne przekraczanie własnego poziomu zastrzeżonego zawsze może zostać skorygowane w kolejnych rundach negocjacyjnych.

Ostatnim istotnym elementem badania miała być weryfikacja rzekomej przewidywalności wyniku negocjacji dystrybucyjnych w oparciu o oferty otwarcia negocjatorów. Bowiem oprócz znanego powszechnie efektu zakotwiczenia [1] istnieje również przekonanie, iż negocjacje dystrybucyjne kończą się na ogół kompromisem leżącym gdzieś w połowie między ofertami otwarcia obydwu stron [10]. Tymczasem z przeprowadzonego eksperymentu wynika, iż wartości kontraktu wynegocjowanego przez strony (o wartości średniej 4062,13 zł) różniły się od teoretycznego* średnio o 504,80 zł, przy odchyleniu standardowym tej średniej wynoszącej 496,45 zł. Trudno zatem uznać ten wynik, bardzo zróżnicowany, za potwierdzający teorię Raiffy, Richardsona i Metcalfa o przewidywalności rezultatów negocjacji dystrybucyjnych na podstawie ofert otwarcia.

3. Alternatywna metoda wspomagania negocjacji dystrybucyjnych z wątkami integracyjnymi

Wyniki opisanego w poprzednim rozdziale eksperymentu skłoniły autora pracy do wstępnej konceptualizacji procedury wspomagania procesu negocjacji dystrybucyjnych z wątkami integracyjnymi (negocjacji typu mieszanego). Wyniki te wyraźnie wskazują na częste zachowania integracyjne wykazujące tendencje do rozszerzania problemu negocjacyjnego przy konsekwentnym operowaniu ceną jako jedyną, formalnie obowiązującą kwestią negocjacyjną. Takie podejście do problemu sprawia, że przedmiot negocjacji nabiera charakteru wielokryterialnego, a jego cena staje się pewnym kryterium globalnym, tj. ekwiwalentem oceny skalarnej wielokryterialnego problemu decyzyjnego. W związku z tym, aby zapewnić porównywalność ofert na poszczególnych etapach negocjacji należałoby zastosować pewną metodę dekompozycyjną,

* Założonego przez Raiffę jako średnia z ofert otwarcia.

pozwalającą na wyznaczenie wartości ocen poszczególnych elementów składowych nowego problemu negocjacyjnego i ewentualnym porównaniu dekomponowanej ceny wyjściowego przedmiotu negocjacji (lub wybranych elementów składowych problemu negocjacyjnego). Niżej przedstawiona zostanie autorska metoda pozwalająca na dokonanie takiej analizy na podstawie metody even swaps.

3.1. Metoda even swaps we wspomaganiu negocjacji integracyjnych

Wywodząc z oryginalnej procedury even swaps [5], będącej metodą wielokryterialnego podejmowania decyzji, autor pracy opracował wcześniej metodę wspomagania negocjacji integracyjnych w zakresie budowy systemu oceny ofert negocjacyjnych [14]. Metoda ta wspomaga negocjatorów w fazie prenegocjacyjnej w analizie preferencji, a następnie wartościowaniu ofert negocjacyjnych, które negocjator planuje uwzględnić w swojej strategii. Sprowadzając ocenę jakości (poziomu satysfakcji z) wielokryterialnej oferty do oceny skalarnej, wyrażonej w jednostce wybranej kwestii negocjacyjnej (tej, którą negocjator uważa za najłatwiejszą w interpretacji), ułatwia ona porównanie ofert oraz analizę różnic jakości między nimi. Na etapie negocjacji właściwych taki system oceny ofert pozwala szybko i jednoznacznie porównać przedkładane przez partnera oferty, analizować skalę dokonywanych ustępstw oraz mierzyć postępy negocjacji [16].

Metoda even swaps bazuje na koncepcji wektorowej dominacji, zgodnie z którą przyjmuje się, że warianty zdominowane są wykluczane z analizy problemu decyzyjnego, gdyż każdy racjonalny decydent wybierze zamiast nich warianty je dominujące (globalnie lepsze). Procedura metody zmierza zatem do eliminowania wszystkich wariantów zdominowanych, a w przypadku braku takowych, do odpowiedniego modyfikowania konsekwencji wariantów decyzyjnych tak, aby były one jednakowe w świetle wybranego kryterium oceny. Wówczas bowiem kryterium takie uznaje się za nieróżnicujące i może zostać usunięte z analizy decyzyjnej. Taka modyfikacja konsekwencji wymaga określenia dla każdego wariantu podstawowej zmiany koniecznej do zniesienia rozpatrywanego kryterium oceny, a następnie wyceny tej zmiany w jednostkach innego kryterium (określenia wartości ekwiwalentnej stanowiącej trade-off).

Rozpatrzmy pewien przykład, w którym ocenie podlegają trzy warianty decyzyjne w negocjacjach między nabywcą a sprzedającymi oferty samochodowe. Ocena dokonywana jest w ramach kryterium ceny i dodatkowego kompletu opon dołączonych do sprzedawanego samochodu (tabela 1).

Tabela 1

Problemu wyboru ofert dealerskich

Oferta	Kryteria oceny samochodu Fiat Seicento	
	cena (PLN)	komplet dodatkowych opon
Dealer 1	5 600	4 opony Dębica na felgach stalowych
Dealer 2	6 000	4 opony Fulda na felgach aluminiowych
Dealer 3	5 000	brak

Decydent przeprowadza analizę even swaps eliminując kryterium dodatkowych opon. Identyfikuje najkorzystniejsze dla niego rozwiązanie w tej kwestii, tj. 4 opony Fulda na felgach aluminiowych, a następnie stara się wyznaczyć różnicę (jakości) między tym rozwiązaniem, a pozostałymi opcjami tej kwestii. Różnica ta musi być wartością ekwiwalentną wyrażoną w jednostkach innej kwestii (tutaj ceny). W definiowaniu tych różnic można posłużyć się standardowymi metodami kosztorysowania lub uwzględnić subiektywną wartość poszczególnych opcji dla decydenta. Następnie dla każdego wariantu poprawiona zostaje jakość pozostałych opcji zdefiniowanych dla kwestii „komplet dodatkowych opon” przy jednoczesnym pogorszeniu wartości opcji kwestii ceny o zdefiniowaną wartość ekwiwalentną (dokonuje się równoważnej wymiany – even swap – jakości obydwu kwestii, tj. poprawy wartości jednej i pogorszenia, o tę wartość równoważną, wartości drugiej kwestii). Przyjmijmy, że decydent wycenił różnicę między opcją „4 opony Dębica...” a „4 opony Fulda...” na 200 zł (nie jest zainteresowany felgami aluminiowymi, 200 zł to ekwiwalent jakości samej opony Fulda wobec Dębica), a różnica „4 opon Fulda...” wobec braku jakichkolwiek opon to 850 zł. Wówczas analiza swaps sprowadza się do wyznaczenia nowych wartości cen w oparciu o wartości ekwiwalentne (tabela 2).

Tabela 2

Analiza even swaps dla problemu wyboru ofert dealerskich

Oferta	Kryteria oceny samochodu Fiat Seicento	
	cena (PLN)	komplet dodatkowych opon
Dealer 1	5 600 5 800	4 opony Dębica na felgach stalowych 4 opony Fulda na felgach aluminiowych
Dealer 2	6 000	4 opony Fulda na felgach aluminiowych
Dealer 3	5 000 5 850	Brak 4 opony Fulda na felgach aluminiowych

Jakość opcji dla kryterium „Komplet dodatkowych opon” została podniesiona kosztem pogorszenia wartości opcji kwestii „Cena”. Decydent może obecnie wyeliminować z analizy decyzyjnej kryterium opon i posługiwać się jedynie skalarnym kryterium ceny, będącym obecnie kryterium zagregowanym, uwzględniającym koszt samochodu doprowadzonego do tzw. postaci wariantu idealnego, tj. realizującego pozostałe kwestie na najlepszych (najbardziej preferowanych) poziomach. Widać zatem, iż kierując się kryterium minimalizacji ceny skumulowanej wybierze wariant dealera 1. W przypadku większej ilości kwestii taką analizę, polegającą na kumulacji wartości ekwiwalentnych na opcjach ceny należałoby przeprowadzić dla wszystkich pozostałych kwestii.

3.2. Dekumulacyjna metoda even swaps we wspomaganii negocjacji dystrybucyjnych

Ideę metody even swaps zaprezentowaną wyżej można, po stosownych modyfikacjach, wykorzystać do wspomagania negocjacji dystrybucyjnych, w których pojawiają się wątki integracyjne. Nowa procedura scoringowa uwzględniać musi jednak całkiem inny mechanizm wyceny przedmiotu negocjacji, jaki będzie miał miejsce w negocjacjach dystrybucyjnych. Odwrotnie niż w algorytmie even swaps dla problemów integracyjnych przyjąć należy obecnie, iż na każdym etapie negocjacyjnym i dla każdej oferty znana jest cena globalna (zagregowana) przedmiotu negocjacji. Przedmiot ten, na skutek zabiegów integracyjnych rozszerzających zakres negocjacji, jest jednak opisany przez pewien kompleks kryteriów (zawierających również kryterium wyjściowe), a zadaniem procedury jest zidentyfikować wartości ekwiwalentne nowych (integracyjnych) elementów problemu negocjacyjnego oraz wykonać swaps tych wartości, mający na celu zredukowanie ceny globalnej i wyznaczenie „ceny netto” wyjściowego przedmiotu negocjacji dystrybucyjnych. Algorytm even swaps rekomendowany do negocjacyjnych problemów dystrybucyjnych będzie więc miał pewien dekulacyjny (subtrakcyjny) charakter, aczkolwiek cała idea operowania wartościami ekwiwalentnymi pozostaje taka sama, jak w wyjściowej procedurze even swaps.

Powyższą ideę dekulacyjnej metody even swaps ilustruje poniższy przykład. Prześledźmy scenariusz pewnych negocjacji.

Nabywca samochodu negocjuje ze sprzedawcą cenę przedmiotu transakcji – samochodu Fiat Seicento. W wyjściowej propozycji sprzedawca żądał ceny 6500 zł. Po wstępnych rozmowach i argumentacji stanu technicznego pojazdu przez nabywcę sprzedawca obniżył cenę do 6000 zł. Nabywca uważał jednak, iż ustępstwo to nie odpowiada potencjalnym kosztom naprawy usterek

zauważonych w pojeździe, dlatego też sprzedawca – zamiast dalej obniżyć cenę – zaproponował gratis kartę stałego klienta, uprawniającą nabywcę do dwóch darmowych przeglądów technicznych samochodu w stacji serwisowej należącej do sprzedawcy (według cennika sprzedawcy każdy z przeglądów kosztuje 150 zł). Nabywca odrzucił propozycję sprzedawcy, blefując, iż sam zna się na mechanice pojazdowej i nie potrzebuje usług zewnętrznych w tym zakresie. Dlatego też sprzedawca zaproponował, iż w zamian może dołożyć do samochodu komplet nowych, zimowych opon Fulda na stalowych felgach (wartość cennikowa według sprzedawcy to 850 zł) za symboliczną cenę 150 zł. Wobec dalszego wahania nabywcy sprzedawca zaproponował (już jako ofertę odejścia) wspomniany wyżej komplet opon gratis plus jeden darmowy przegląd w swym zakładzie serwisowym.

Jak łatwo zauważyć, w powyższym opisie przypadku negocjacyjnego pojawiło się kilka ofert zawierających integracyjne propozycje rozszerzające wyjściowy problem negocjacyjny. Stanowią one typowy zabieg marketingowy stosowany obecnie powszechnie wobec klientów w negocjacjach handlowych (kontraktu na dobra i usługi) bez względu na branżę transakcji. W takiej sytuacji pojawia się więc problem wyceny rzeczywistego, wyjściowego przedmiotu negocjacji (przedmiotu negocjacji dystrybucyjnych). Spróbujemy więc w tym celu zastosować dekulacyjną metodę even swaps.

Historię ofert negocjacyjnych złożonych przez sprzedawcę w toku negocjacji przedstawiono w sposób syntetyczny w tabeli 3.

Tabela 3

Oferty dealerskie w negocjacjach samochodowych

Oferta	Kryterium dystrybucyjne	Kryteria integracyjne proponowane przez dealera	
	cena (PLN)	komplet dodatkowych opon	przeglądy techniczne
1	6 500	brak	brak
2	6 000	brak	brak
3	6 000	brak	2
4	6 150	Fulda + felgi stalowe	brak
5	6 000	Fulda + felgi stalowe	1

Nabywca może porównać jakość powyższych ofert stosując odwróconą procedurę even swaps, dokonującą subtrakcji wartości ekwiwalentnych na kryterium dystrybucyjnym. W pierwszym kroku musi on zdefiniować wartości sum ekwiwalentnych (wyceny) dla każdej opcji kryteriów integracyjnych. Przyjmijmy, że dla kwestii opon opcja „Fulda z felgami” jest dla niego warta 450 zł

(nabywca jest indyferentny wobec felg, gdyż posiada własne zapasowe felgi, a wartość ekwiwalentna obejmuje same opony). Dla kwestii przeglądów technicznych każdy przegląd nabywca wycenia na 60 zł (które musiałby zapłacić u zaprzyjaźnionego serwisanta). Uwzględniając fakt, iż drugi z przeglądów byłby realizowany dopiero za rok i biorąc pod uwagę wartość pieniądza w czasie wartość ekwiwalentna 2 przeglądów wyniesie 123,60 zł. Znając wartości ekwiwalentne można obecnie przeprowadzić analizę swaps sprowadzając wszystkie wartości opcji do poziomu „brak”, a przez to wyznaczyć cenę netto dla wyjściowego problemu negocjacji dystrybucyjnych (negocjowania ceny samochodu). Analizę dekulacyjną zmodyfikowanej, subtrakcyjnej metody even swaps przedstawiono w tabeli 4.

Tabela 4

Analiza dekulacyjna even swaps dla ofert dealerskich
w negocjacjach samochodowych

Oferta	Kryterium dystrybucyjne	Kryteria integracyjne proponowane przez dealera	
	cena (PLN)	komplet dodatkowych opon	przeglądy techniczne
1	6 500	brak	brak
2	6 000	brak	brak
3	6 000 5 876,40	brak	2 brak
4	6 150 5 700	Fulda + felgi stalowe brak	brak
5	6 000 5 490	Fulda + felgi stalowe brak	1 brak

Z przeprowadzonej analizy even swaps wynikają zarówno dokładna ścieżka ustępstw sprzedawcy oraz rzeczywiste wartości negocjowanego pojazdu (a nie kompleksu przedmiotów negocjacji, którymi sam nabywca nie musiał być nawet zainteresowany). Nabywca jest więc w stanie obecnie jednoznacznie zdecydować, czy i która z propozycji sprzedawcy spełnia jego poziom rezerwacji b . Jeśli na liście ofert sprzedawcy znajduje się kilka propozycji spełniających poziom zastrzeżony, wówczas nabywca może porównać je dodatkowo z poziomem aspiracji (czy też pewnym planowanym do ujawnienia poziomem b'), aby sprawdzić, czy w danej rundzie negocjacji (uwzględniając potencjalnie negatywne efekty presji czasu) proponowana oferta jest dla niego korzystna i może zostać zaakceptowana jako satysfakcjonujący kompromis.

Podsumowanie

Rozważania przedstawione w niniejszej pracy miały charakter zarówno systematyzujący, jak i eksploracyjny oraz koncepcyjny. Przedstawiony opis negocjacji dystrybucyjnych miał na celu przybliżenie klasycznego ujęcia negocjacji nad podziałem ograniczonego dobra (pojedynczej kwestii negocjacyjnej). Zaplanowane i opisane badanie miało na celu weryfikację rzeczywistych zachowań uczestników negocjacji dystrybucyjnych i wskazało na pojawianie się w tych negocjacjach wątków integracyjnych, otwierających możliwości alternatywnego podejścia do modelowania i wspomagania takiego procesu. O możliwości istnienia takiego zjawiska w negocjacjach dystrybucyjnych pisano już zresztą wcześniej w literaturze, definiując taki typ negocjacji, obok negocjacji dystrybucyjnych i integracyjnych, jako negocjacje typu mixed [6]. Badanie potwierdziło jednak, że zjawisko zachowań integracyjnych w negocjacjach dystrybucyjnych wcale nie jest rzadkie, stąd też zasadność podejmowania pewnych badań i analiz zmierzających do wspomagania tego typu negocjacji. Chcąc jednakże uzyskać bardziej wiarygodne i statystycznie potwierdzone wyniki, planuje się powtórzenie takiego badania w przyszłości, dodatkowo rozszerzając go o zagadnienia pozwalające na weryfikację motywów takiego, a nie innego działania negocjatorów (np. zawyżania wartości poziomów zastrzeżonych, wprowadzania wątków integracyjnych, interpretacji wyniku jako ceny oferty integracyjnej itp.).

Opierając się na wynikach eksperymentu w pracy starano się nakreślić pewną możliwość wspomagania tego typu negocjacji w zakresie wartościowania pojawiających się w toku negocjacji ofert integracyjnych. Wartościowanie to zdecydowano się prowadzić opierając się na prostej i intuicyjnej metodzie even swaps, która została już wcześniej wykorzystana przez autora pracy do wspomagania negocjacji typu integracyjnego, a ponadto dała się łatwo zaimplementować w systemie wspomagania negocjacji automatyzującym i ułatwiającym przygotowawcze prace negocjacyjne [15]. Dla zaproponowanego podejścia przewiduje się również takowa implementację w przyszłości, która jednakże poprzedzona musi zostać pracami analitycznymi zmierzającymi do formalizacji i algorytmizacji metody oraz utworzenia dla niej specjalnego protokołu negocjacyjnego, według którego (a zgodnie z przyjętą metodologią) negocjatorzy prowadzić będą zarówno prace prenegocjacyjne, jak i negocjacje właściwe.

Ciekawą konkluzją płynącą zarówno z wyników eksperymentu, jak i zaproponowanej metody wspomagania negocjacji dystrybucyjnych z wątkami integracyjnymi jest możliwość (lub potencjalna konieczność) prowadzenia

analiz postoptymalizacyjnych dla uzyskanego w tego typu negocjacjach kompromisu. Pojawianie się dodatkowych kryteriów integracyjnych (przedmiotów rozszerzających zakres negocjacji) rodzi pytanie o to, czy rzeczywiście jedynym wyznacznikiem jakości kompromisu ma być cena netto (cena wyjściowego przedmiotu negocjacji) czy też nie zastosować analizy wielokryterialnej uwzględniającej włączenie do oceny kompromisu przynajmniej części zaproponowanych kwestii integracyjnych (tych na których zależy nabywcy/sprzedawcy). Jeśli takie będzie oczekiwanie negocjatorów, wówczas pojawi się w tych negocjacjach problem typowy i znany z negocjacji czysto integracyjnych, czyli problem rzutowania kompromisów na przestrzeń wyników obydwu graczy i poszukiwania wariantów efektywnych i dominujących uzyskany przez negocjatorów wynik.

Literatura

1. Bazerman M.H., Neale M.A. (1997). *Negocjując racjonalnie*. Polskie Towarzystwo Psychologiczne, Olsztyn.
2. Cenker E.M. (2002). *Negocjacje*. Wydawnictwo Wyższej Szkoły Komunikacji i Zarządzania w Poznaniu, Poznań.
3. Chatterjee K., Samuelson W. (1981). *Simple Economics of Bargaining*. Distribution Paper. Boston University.
4. Fisher R., Ury W. (1995). *Dochodząc do tak – negocjowanie bez poddawania się*. Państwowe Wydawnictwo Ekonomiczne, Warszawa.
5. Hammond L., Keeney R., Raiffa H. (1998). *Even Swaps: A Rational Method for Making Trade-offs*. Harvard Business Review (march).
6. Kersten G.E. (2001). *Modeling Distributive and Integrative Negotiations. Re-view and Revised Characterization*, Group Decision and Negotiation, 10, s. 493-514.
7. *Proceedings of the Group Decision and Negotiation'2007 Conference*. (2007). Red. G.E. Kersten, J. Rios, E. Chen. InterNeg Research Centre, Montreal.
8. Lewicki R., Saunders D., Minton J. (1999). *Negotiation*. The MacGraw-Hill Companies Inc., Boston.
9. *Negotiation Analysis*. Red. H.P. Young. The University of Michigan Press.
10. Raiffa H. (1982). *The Art and Science of Negotiation*. Harvard University Press, Cambridge.
11. Raiffa H., Richardson J., Metcalfe D. (2002). *Negotiation Analysis*. The Belknap Press of Harvard University Press, Cambridge.
12. Roszkowska E. (2011). *Wybrane modele negocjacji*. Wydawnictwo Uniwersytetu w Białymstoku, Białystok.

13. Thompson L. (1998). *The Mind and Heart of the Negotiator*. Prentice Hall, Upper Saddle River, New Jersey.
14. Wachowicz T. (2008). Metoda Even Swaps w tworzeniu systemu oceny ofert negocjacyjnych. W: *Metody i Zastosowania Badań Operacyjnych '07*. Red. D. Kopańska-Bródka. AE, Katowice, s. 297-308.
15. Wachowicz T. (2008). NegoCalc: Spreadsheet Based Negotiation Support Tool with Even-Swap Analysis. In: *Group Decision and Negotiation 2008: Proceedings – Full Papers*. Eds. J. Climaco, G. Kersten, J.P. Costa. INESC Coimbra, s. 323-329.
16. Wachowicz T. (2010). Metody i narzędzia wspomagania fazy prenegocjacyjnej. *Decyzje*, 14, s. 55-84.
17. Walton R.E., McKersie R.B. (1965). *A Behavioral Theory of Labor Negotiations*. McGraw-Hill, New York.
18. Watson J. (2004). *Strategia. Wprowadzenie do teorii gier*. Wydawnictwa Naukowo-Techniczne, Warszawa.

DISTRIBUTIVE NEGOTIATIONS AND THE NEGOTIATORS' INTEGRATIVE ACTIONS. AN ANALYSIS AND SUPPORT

Summary

In this work we focus on the distributive negotiations in which the single negotiation issue is taken into consideration only. We present the different approaches to describing, modeling, analyzing and supporting this kind of negotiations and compare them with the integrative negotiation, where the multitude of issues is discussed. Then we present and analyze the results of the in-class negotiation experiment. Deriving from these result we propose the novel approach to these distributive negotiations in which the integrative moves are undertaken by the negotiators.

ZARZĄDZANIE PROJEKTAMI

Marek Janczura

Politechnika Wrocławska

Dorota Kuchta

Politechnika Wrocławska

Wyższa Szkoła Oficerska Wojsk Lądowych im. gen. Tadeusza Kościuszki we Wrocławiu

HARMONOGRAMOWANIE REAKTYWNE

A RYZYKO W PROJEKCIE

Wprowadzenie

Jednym z kluczowych elementów projektu są jego zasoby ludzkie. To od człowieka w znacznej mierze zależy powodzenie realizacji zadań w czasie zgodnym z tym przewidzianym w harmonogramie. Ludzie są cały czas podatni na wiele niespodziewanych zdarzeń. Przez to zarządzanie zasobami powinno rozpocząć się już na etapie planowania, a także być kontynuowane w trakcie realizacji projektu.

Wyróżniamy dwa typy ryzyka związanego z zasobami ludzkimi. Pierwsze z nich to ryzyko negatywne, które oznacza, że ludzie pracują wolniej niż zakładaliśmy lub też stają się niedostępni dla projektu. Ryzyko pozytywne ma miejsce wtedy, gdy ludzie pracują szybciej niż szacowaliśmy lub też pojawiają się dodatkowe i niespodziewane wcześniej zasoby do wykorzystania.

W projekcie problem niepewności dotyczy nie tylko ludzi, ale również pozostałych zasobów i zdarzeń mających wpływ na czas trwania zadań. Zadania nowe, czyli te realizowane po raz pierwszy, stanowią grupę zadań trudnych do estymacji ze względu na brak wiedzy historycznej. Jednocześnie w trakcie realizacji projektu pojawiają się nieprzewidziane zdarzenia, które dodatkową zwiększają niepewność czasu realizacji.

Wyróżnia się dwa podejścia budowy harmonogramu [8]. Pierwsze z nich – podejście proaktywne – zakłada stworzenie na etapie planowania harmonogramu odpornego na nieprzewidziane zdarzenia, mogące się pojawić w trakcie realizacji projektu. Podejście reaktywne zakłada aktualizację harmonogramu oraz podjęcie akcji naprawczych w trakcie realizacji projektu. W obu tych podejściach, w celu lepszej kontroli niepewności czasów trwania zadań, można

zastosować liczby rozmyte do estymacji długości zadań. Jediną pracą traktującą o wykorzystaniu liczb rozmytych do budowy i pracy z harmonogramem rozmytym w modelu reaktywnym jest praca Janczury i Kuchty [5]. Koncepcję wykorzystania liczb rozmytych w podejściu reaktywnym opisuje również Long i Ohsato [4]. Wychodzą oni jednak od konstrukcji harmonogramu deterministycznego, gdzie rozmyte czasy trwania zadań wykorzystują do jego rozszerzenia o dodatkowy bufor czasowy.

W niniejszym opracowaniu prezentujemy rozszerzony model harmonogramowania projektu. W poprzedniej pracy [5] przedstawiliśmy model z rozmytymi czasami trwania, w odpowiedzi na problem radzenia sobie z niepewnością estymacji długości zadań. Stworzyliśmy również mechanizm naprawy harmonogramu polegający na aktualizowaniu rozmytych czasów trwania w trakcie realizacji projektu.

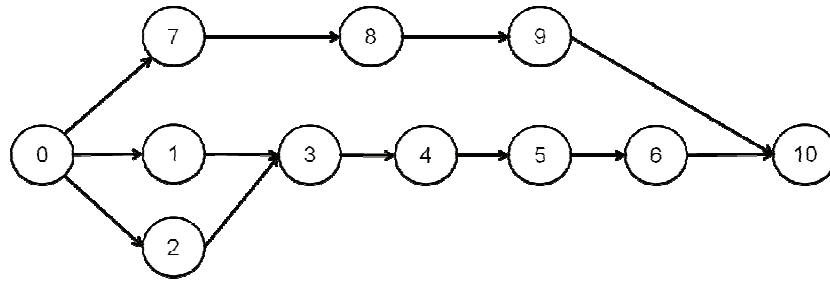
Teraz proponujemy rozszerzenie modelu o uwzględnienie ryzyka związanego z zasobami. Prezentujemy możliwość wyboru różnych strategii przydzielania zasobów, w zależności od stopnia pesymizmu/optymizmu w ocenie czasu trwania poszczególnych zadań, co pozwala na dopasowanie harmonogramowania poszczególnych zadań do preferencji kierownika projektu.

Najpierw przedstawimy podstawowe definicje, a następnie proponowany algorytm, który będzie zilustrowany przykładem liczbowym. W kolejnych rozdziałach będziemy się posługiwać terminem zasoby w kontekście zasobów ludzkich w projekcie.

1. Tło zagadnienia

1.1. Sieć projektu

Do opisu projektu przyjmijmy następujące oznaczenia: $S = \langle V, A, T \rangle$, gdzie V jest skończonym zbiorem węzłów (zdarzeń), $A \subset V \times V$ jest skończonym zbiorem krawędzi (zadań), a T jest zbiorem estymowanych czasów trwania. W niniejszej pracy zakładamy, że sieć projektu jest reprezentowana przez model AOA (Activity On Arc) rys. 1. Zadania są reprezentowane przez uporządkowaną parę węzłów $(i, j) \in A$. Zakładamy, że zadania są realizowane tak szybko jak to możliwe oraz że zachodzą między nimi relacje poprzedzania (koniec-początek). Ponadto rozpatrujemy zadania o ustalonej liczbie pracy (np. o ustalonej liczbie roboczości, jaka musi być wykonana – długość takiego zadania będzie się zmieniać w zależności od liczby przydzielonych zasobów). Elementy zbioru T definiujemy jako t_{ij} – czas trwania zadania (i, j) .



Rys. 1. Przykład sieci projektu

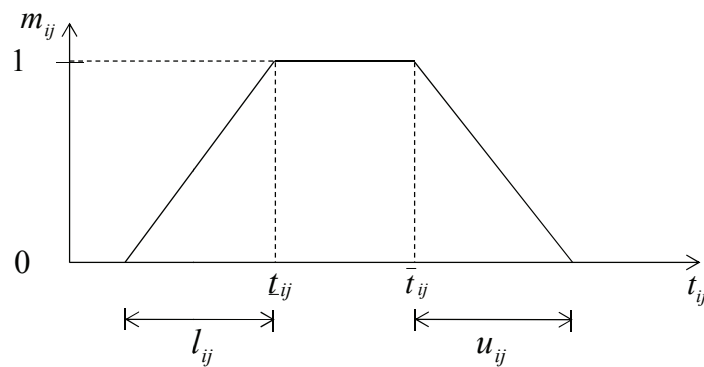
2. Liczby rozmyte jako miara niepewności

2.1. Rozmyty czas trwania zadań

Do opisu czasu trwania zadań wykorzystane zostaną liczby rozmyte [9]. Niech $\tilde{T}_{ij}$ będzie rozmytą liczbą reprezentującą czas trwania zadania (i, j) , o następującej definicji

$$\tilde{T}_{ij} = (\underline{t}_{ij}, \bar{t}_{ij}, l_{ij}, u_{ij}) \quad (1)$$

gdzie $\underline{t}_{ij}$ i $\bar{t}_{ij}$ są odpowiednio dolnym i górnym ograniczeniem przedziału czasu z najwyższą szansą realizacji. Dwa kolejne parametry l_{ij} i u_{ij} (rys. 2) definiują długości przedziałów ze zmienną szansą realizacji. Są one również oceną popełnionego błędu przy estymacjach. Wykorzystywane liczby rozmyte oparte są na modelu trapezowym [7] (rys. 2).

Rys. 2. Rozmyty czas trwania zadania (i, j)

Przedziały są dobrane na podstawie wywiadu z ekspertem i uwzględniają niepewność dotyczącą czasu trwania zadania.

Oznaczmy sieć projektu przez $\tilde{S} = \langle V, A, \tilde{T} \rangle [3]$, gdzie V i A są zdefiniowane jak w 1.1, natomiast czasy trwania zadań będą teraz reprezentowane przez liczbę rozmytą $\tilde{T}$.

Rysunek 2 przedstawia funkcję przynależności trapezowej liczby rozmytej (1). Wzór na tę funkcję jest następujący

$$m_{ij}(t) = \begin{cases} 0, & t < \underline{t}_{ij} - l_{ij} \\ \frac{l_{ij} - \underline{t}_{ij} + t}{l_{ij}}, & t \in [\underline{t}_{ij} - l_{ij}, \underline{t}_{ij}] \\ 1, & t \in [\underline{t}_{ij}, \bar{t}_{ij}] \\ \frac{u_{ij} - t + \bar{t}_{ij}}{u_{ij}}, & t \in (\bar{t}_{ij}, \bar{t}_{ij} + u_{ij}] \\ 0, & t > \bar{t}_{ij} + u_{ij} \end{cases} \quad (2)$$

Funkcja przynależności (2) w rozpatrywanym przypadku definiuje stopień możliwości realizacji zadania (i, j) w czasie t .

2.2. Rozmyta ścieżka krytyczna

Ścieżka krytyczna jest najdłuższą sekwencją zadań realizowanych w projekcie. Wyznacza ona również całkowity czas trwania projektu. W niniejszym opracowaniu proponujemy następujący algorytm obliczenia ścieżki krytycznej:

- 1) obliczenie rozmytych czasów trwania zadań,
- 2) obliczenie deterministycznych oszacowań czasów trwania zadań na podstawie wybranej strategii przydziału zasobów oraz dostępności zasobów,
- 3) wybranie ścieżki krytycznej jako najdłuższej sekwencji zadań w projekcie.

Deterministyczne oszacowanie czasu trwania polega na zastąpieniu liczby rozmytej pewną liczbą rzeczywistą. Można to robić w różny sposób, ale zawsze zależy to od preferencji decydenta. Szczegóły będą omówione w podrozdziale 3.1.

3. Ograniczenia zasobów ludzkich

Przyjmijmy następujące oznaczenie związane z liczbą zasobów w projekcie ($t = 1, \dots, H$, gdzie H oznacza maksymalny horyzont czasu, w którym może być realizowany projekt. Przyjmujemy, że podstawową jednostką planowania jest dzień):

- npp^t – liczba zasobów dostępnych w projekcie w dniu t ,
- npd^t – liczba zasobów wymaganych w projekcie w dniu t ,
- npt_{ij}^d – liczba zasobów wymaganych w dniu d zadania (i, j) .

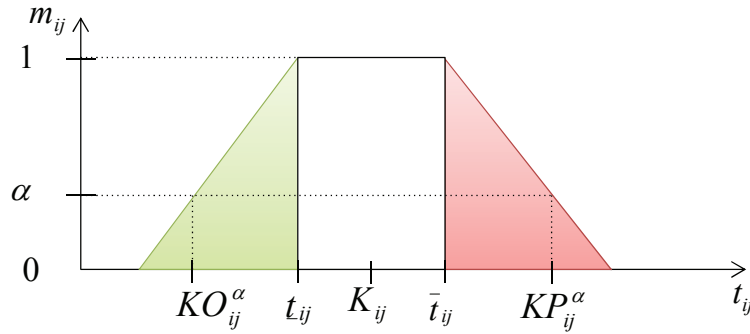
Zakładamy, że dostępność zasobów może się zmieniać w trakcie realizacji projektu. Dzięki temu możliwe będzie uwzględnienie w modelu zmniejszenia czy zwiększenia liczby zasobów w danym dniu (choroba, zatrudnienie dodatkowe). Stąd wynika potrzeba monitorowania zapotrzebowania oraz dostępności zasobów wielokrotnie po rozpoczęciu projektu.

W niniejszej pracy zakładamy nierozróżnialność (jeden typ) zasobów. Zadania mogą mieć również zmienne zapotrzebowanie na zasoby (dziennie). Rozwiązanie to pozwala zarządzać zasobami na poziomie poszczególnych zadań. Proponujemy następujący algorytm przydziału zasobów do zadań:

- 1) obliczenie rozmytego czasu trwania dla danego zadania (i, j) uwzględniającego niepewność początkową (patrz 2.1),
- 2) harmonogramowanie zadań (patrz 3.2),
- 3) wybór strategii przydziału zasobów (patrz 3.1),
- 4) aktualizacja czasu trwania zadania na podstawie wybranej strategii i dostępności zasobów (patrz 3.3),
- 5) dobór odpowiedniej ilości zasobów do każdego dnia czasu trwania zadania.

3.1. Strategia optymisty/pesymisty

Zasoby ludzkie są przydzielane w zależności od przyjętej strategii kierownika projektu. Polega ona na wyborze czasu, na jaki powinny być zarezerwowane zasoby w danym zadaniu. Wybór strategii może być dokonany osobno na poziomie każdego zadania. Możemy też założyć, iż przydział zasobów dla wszystkich zadań jest zgodny z jedną strategią. W trakcie realizacji projektu, możemy dokonać zmiany strategii poszczególnych zadań według preferencji eksperta.



Rys. 3. Możliwe wartości deterministycznego czasu trwania zadania (i, j) dla strategii uśrednionej, K_{ij} optymistycznej KO_{ij}^α oraz pesymistycznej KP_{ij}^α

Punktem wyjścia są rozmyte czasy trwania zadań. W zależności od preferencji kierownik projektu może wybrać jedną z trzech strategii: bazującą na średniej oraz dwie strategie polegające na optymistycznym lub pesymistycznym ocenie przez eksperta czasu trwania zadania (rys. 3). Rezerwacja zasobów w tych zadaniach odbywa się w następujący sposób:

1. Strategia Uśredniona – rezerwacja zasobów na okres K_{ij} – jest to wartość średnia liczby rozmytej, która może być wyznaczona np. za pomocą tzw. indeksu Yagera [2].

2. Strategia Optymisty – rezerwacja zasobów na okres KO_{ij}^α , gdzie

$$KO_{ij}^\alpha = \alpha l_{ij} + t_{ij} - l_{ij} \quad (3)$$

3. Strategia Pesymisty – rezerwacja zasobów na okres KP_{ij}^α , gdzie

$$KP_{ij}^\alpha = \bar{t}_{ij} + u_{ij} - \alpha u_{ij} \quad (4)$$

Wartości KO_{ij}^α oraz KP_{ij}^α są obliczone na podstawie funkcji przynależności [2]. Parametr α przyjmuje wartości od 0 do 1 i służy stopniowaniu podejścia optymistycznego oraz pesymistycznego.

3.2. Harmonogramowanie zadań

Kluczową rolę przy podziale zasobów odgrywają priorytety zadań. Zasoby są przydzielane do zadań, które nie mają poprzedników nieuwzględnionych jeszcze w harmonogramie, przy czym wśród tych zadań o kolejności decyduje priorytet. Priorytety zadań są tutaj obliczane na podstawie liczby ich następników

$N_{nas_{ij}}$ – liczba następników zadania (i, j)

Uzasadnieniem takiego wyboru sposobu liczenia priorytetu zadania jest fakt, że duża liczba następników potencjalnie oznacza znaczną liczbę problemów w przyszłości, dlatego zadania o dużej liczbie następników dobrze jest zrealizować jak najwcześniej. Niemniej jednak dopuszczalne są również inne sposoby wyznaczania priorytetu poszczególnych zadań, również eksperckie.

Najwyższy priorytet ma to zadanie, które ma najwięcej następników wśród wszystkich zadań projektu (porównaj z [6]). Zakładamy, że im niższy wskaźnik priorytetu, tym wyższy priorytet zadania. Zatem zadanie o najwyższym prioryecie ma wskaźnik priorytetu równy 1. W sytuacji, gdy dwa zadania mają tę samą liczbę następników, decydujące znaczenie będzie miała kolejność zadań podana przez eksperta (zadanie zdefiniowane wyżej na liście będzie miało wyższy priorytet). Zastosowanie priorytetów rozwiązuje problem kolejności przydzielania zasobów w sytuacji, gdy zadania mogłyby być realizowane w tym samym czasie, ale brakuje na to zasobów. Rozwiązanie problemu niewystarczającej liczby dostępnych zasobów jest przedstawione w rozdziale 3.3.

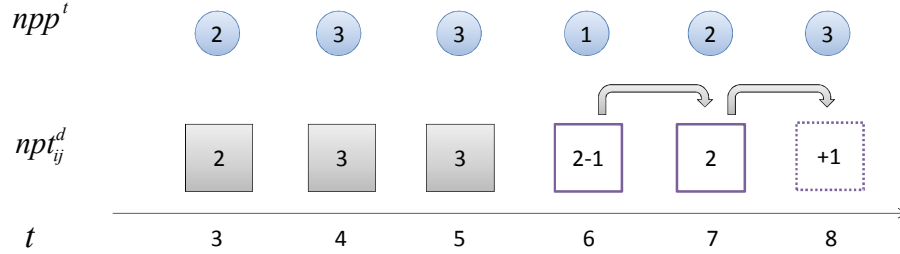
3.3. Rozwiązywanie konfliktu zasobów

W trakcie realizacji projektu (w dniu t , gdzie $t = 1, \dots, H$) może się zdarzyć, że będzie brakować dostępnych zasobów na realizację zadania w konkretnym dniu

$$npt_{ij}^d > npp^t \quad (5)$$

gdzie dzień t jest d -tym dniem realizacji zadania (i, j) , a npp^t jest uaktualnioną liczbą zasobów dostępnych w projekcie w dniu t . W takiej sytuacji proponujemy przeniesienie niewykonanej pracy na dzień kolejny (patrz linia 6, tabela 1).

Może się okazać, iż tego następnego dnia będą wolne zasoby, które nadrobią opóźnioną pracę. W przeciwnym wypadku, zapotrzebowanie na pracę zostanie przekazane na następny dzień. W sytuacji, gdy w trakcie czasu przewidzianego na realizację zadania nie zostanie nadrobiona niewykonana praca, zadanie zostaje wydłużone (rys. 5 oraz linia 11, tabela 1).



Rys. 5. Wydłużenie czasu trwania zadania ze względu na brak dostępnych zasobów

Niech beg_{ij} oznacza czas rozpoczęcia zadania (i, j) oraz założmy, że zadanie (i, j) jest na początku estymowane według strategii uśrednionej – K_{ij} . Algorytm rozwiązywania konfliktu zasobów dla zadania (i, j) wygląda jak w tabeli 1.

Tabela 1

Algorytm przydzielania zasobów dla zadania (i, j)

1: $d = 0$	pierwszy dzień zadania (i, j)
2: FOR $t = beg_{ij} : beg_{ij} + K_{ij}$	dla każdego dnia zadania (i, j)
3: IF $npt_{ij}^d \leq npp^t$ THEN	są dostępne zasoby
4: $npp^t = npp^t - npt_{ij}^d$	zarezerwuj zasoby projektowe
5: ELSE	brak dostępnych zasobów
6: $npt_{ij}^{d+1} = npt_{ij}^{d+1} + (npt_{ij}^d - npp^t)$	zwiększ zapotrzebowanie na zasoby w dniu następnym $d + 1$
7: $npt_{ij}^d = npp^t$	w dniu d ustaw tylko zużyte zasoby
8: $npp^t = 0$	zarezerwuj zasoby projektowe
9: IF $t = beg_{ij} + K_{ij} - 1$ THEN	jeśli ostatni dzień zadania
10: $K_{ij} = K_{ij} + 1$	wydłuż zadanie o jeden dzień
11: $d = d + 1$	przejdź do kolejnego dnia zadania

Zaktualizowany czas trwania zadania (i, j) obliczony na podstawie wybranej strategii oraz informacji o dostępnych zasobach oznaczamy przez KR_{ij} .

4. Harmonogramowanie reaktywne

W tym rozdziale prezentujemy metodę wielokrotnej aktualizacji harmonogramu w trakcie realizacji projektu. Polega ona na monitorowaniu odchyleń w projekcie i mierzeniu ich wpływu na całkowity czas trwania projektu. Głównym rdzeniem algorytmu harmonogramowania reaktywnego jest aktualizacja czasów trwania zadań na podstawie:

- 1) informacje o opóźnieniach i przyspieszeniach w zadaniach,
- 2) ryzyka związanego z zasobami.

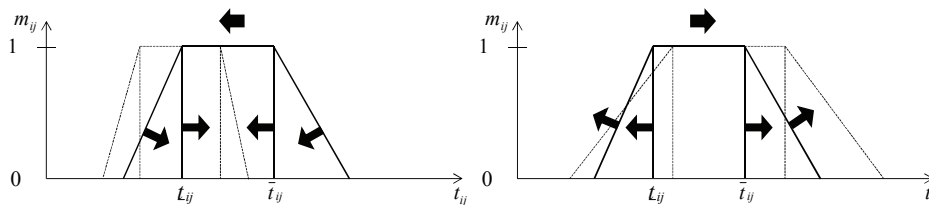
Algorytm aktualizacji polega na generowaniu całego harmonogramu od początku projektu na podstawie informacji uzyskiwanych w trakcie realizacji.

4.1. Aktualizacja liczb rozmytych

Rozmyty czas trwania zadania (i, j) w momencie k ($k = 1$ odpowiada fazie planowania, $k > 1$ fazie realizacji projektu) w trakcie realizacji projektu jest zdefiniowany następująco

$$\tilde{T}_{ij}^k = (\underline{t}_{ij}^k, \bar{t}_{ij}^k, l_{ij}^k, u_{ij}^k) \quad (6)$$

W pracy [1] dla liczby rozmytej [6] zaproponowaliśmy jej aktualizację polegającą na rozszerzaniu przedziałów niepewności w sytuacji opóźnienia realizacji projektu oraz zwężania przedziałów w sytuacji przyspieszenia realizacji (rys. 6).



Rys. 6. Rysunek po prawej (lewej) przedstawia rozszerzenie (zwężanie) liczby rozmytej w odpowiedzi na opóźnienie (przyspieszenie) realizacji zadania (i, j)

4.2. Reagowanie na zmianę dostępności zasobów (dialog z użytkownikiem)

Następnym krokiem aktualizacji czasu trwania zadań jest weryfikacja strategii opisanej w rozdziale 3.1. Na tym etapie, na podstawie informacji z częściowej realizacji, kierownik projektu może zmienić strategię na tę bliższą obecnej sytuacji. Weźmy np. zadanie (i, j) , które było estymowane na początku zgodnie ze strategią uśrednioną. Niech w punkcie kontrolnym $k = 1$ zadanie będzie zrealizowane częściowo z opóźnieniem. Algorytm aktualizacji zadania (i, j) w punkcie kontrolnym $k = 1$ przebiega następująco:

1. Aktualizacja rozmytego czasu trwania zadania (i, j) (patrz 4.1):

- uwzględnia ona opóźnienie zadania (i, j) ,
- jeśli zadanie nadal będzie estymowane zgodnie z tą samą strategią (uśrednioną), to jego uaktualniony planowany czas trwania będzie dłuższy niż czas planowany na początku.

2. Aktualizacja strategii dla zadania (i, j) :

- kierownik projektu może stwierdzić, iż zadanie będzie dalej realizowane z opóźnieniem (ocenia, iż szanse na to są większe niż na początku). Decyduje wtedy na zmianę strategii na pesymistyczną (na danym poziomie α , patrz 3.1). Dzięki temu uwzględni on w planowaniu dalszej części projektu nie tylko aktualne opóźnienie zadania (i, j) , ale również możliwość pojawienia się kolejnych opóźnień w tym zadaniu.

3. Przydział zasobów dla zadania (i, j) :

- zasoby zostają przydzielone do zadania zgodnie z aktualnym zapotrzebowaniem zadania (i, j) , a także ich dostępnością w projekcie (tabela 1),
- dzięki temu na bieżąco uwzględniamy zmiany dostępności zasobów.

Powyższa procedura powtarzana jest dla wszystkich zadań w projekcie w kolejności według ich priorytetów (patrz 3.2). Kierownik projektu, wykorzystując algorytm dla wszystkich zadań, dokonuje ich aktualizacji z uwzględnieniem opóźnień i przyspieszeń w realizacji, a także zmian dostępności zasobów.

4.3. Aktualizacja czasu trwania projektu

Po aktualizacji rozmytych czasów trwania i strategii oraz przydzieleniu zasobów otrzymujemy zaktualizowane deterministyczne oszacowanie czasu trwania projektu. Czas jest obliczany na podstawie wszystkich aktualnych informacji, również opierając się na strategiach wybranych przy szacowaniu

deterministycznych czasów trwania zadań. Dzięki powtarzaniu tej czynności wielokrotnie w trakcie realizacji projektu można na bieżąco śledzić postęp prac w całym projekcie, kontrolować wszystkie zmiany, a także prognozować ich wpływ na przedłużenie całkowitego czasu trwania projektu.

5. Przykład

W tym rozdziale przedstawiamy symulację przykładowego projektu. W pierwszej części (patrz 5.1) omówiona zostanie estymacja czasów trwania zadań, wybór strategii oraz końcowy harmonogram z uwzględnieniem dostępności zasobów ludzkich. Druga część (patrz 5.2) to przykładowa symulacja aktualizacji projektu w czasie realizacji wraz z weryfikacją wybranej strategii.

5.1. Planowanie

Sieć projektu odpowiadająca niniejszemu przykładowi znajduje się na rysunku rys. 1. Relacje pomiędzy zadaniami są przedstawione w tabeli 2.

Tabela 2

Relacje pomiędzy zadaniami w projekcie

Zadanie (i, j)	Poprzednik zadania (i, j)			
(7, 8)				
(8, 9)	(7, 8)			
(1, 3)				
(2, 3)				
(3, 4)	(1, 3)	(2, 3)		
(4, 5)	(1, 3)	(2, 3)	(3, 4)	
(5, 6)	(1, 3)	(2, 3)	(3, 4)	(4, 5)

Priorytety zadań obliczone zgodnie z algorytmem opisanym w 3.2 są podane w tabeli 3.

Tabela 3

Priorytety zadań w projekcie

Zadanie (i, j)	Priorytet
(7, 8)	5
(8, 9)	7
(1, 3)	2
(2, 3)	1
(3, 4)	3
(4, 5)	4
(5, 6)	6

Projekt jest realizowany zgodnie ze strategią uśrednioną przydzielania zasobów. Dla przykładu obliczyliśmy również wartość odpowiadającą strategii optymistycznej na poziomie $\alpha = 1$ oraz strategii pesymistycznej na poziomie $\alpha = 0$.

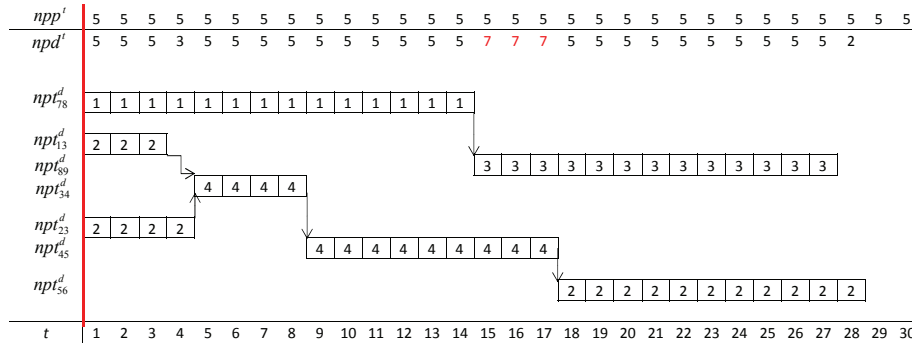
Tabela 4

Zestawienie czasów trwania zadań przed uwzględnieniem zasobów w punkcie kontrolnym projektu $k = 1$

Zadanie (i, j)	$\underline{t}_{ij}^1$	$\bar{t}_{ij}^1$	l_{ij}^1	u_{ij}^1	K_{ij}^1	$KO_{ij}^{1,1}$	$KP_{ij}^{1,0}$
(7, 8)	12	15	2	4	14	12	19
(8, 9)	10	15	3	5	13	10	20
(1, 3)	2	4	1	1	3	2	5
(2, 3)	2	6	1	1	4	2	7
(3, 4)	3	5	2	2	4	3	7
(4, 5)	8	11	3	1	9	8	12
(5, 6)	8	13	2	4	11	8	17

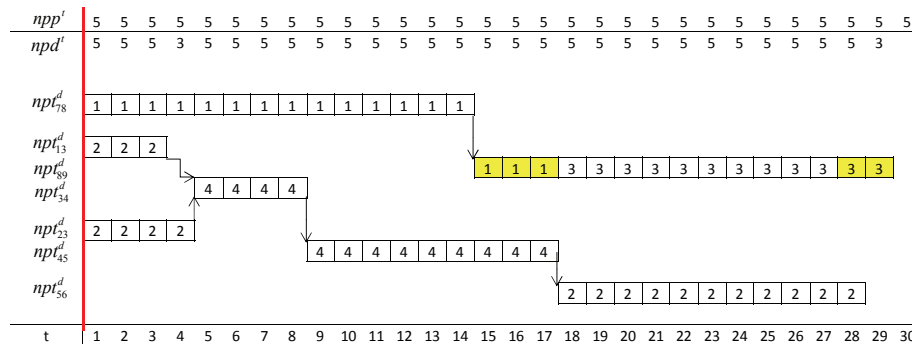
W tabeli 4 zestawiono planowane długości zadań (przed uwzględnieniem zasobów). Rozmyte czasy trwania zadań znajdują się w kolumnach 2-5. Kolejne kolumny przedstawiają czasy trwania zadań dla trzech strategii: wartość średnią, wartość optymistyczną (dla $\alpha = 1$) oraz wartość pesymistyczną (dla $\alpha = 1$). Dla zadania (4, 5) całkowita długość zadania, w zależności od wybranej strategii, może wynosić 8, 9 albo nawet 12 dni. Decyzja, którą wartość wykorzystać do rezerwacji zasobów w projekcie, należy do kierownika projektu.

Na rys. 7 przedstawiono harmonogram po estymacji w punkcie kontrolnym projektu $k = 1$ zgodnie ze strategią uśrednioną. W 15, 16 i 17 dniu projektu, brakuje zasobów na realizację równoległą zadań (8, 9) oraz (4, 5). Ponieważ zadanie (8, 9) ma niższy priorytet (tabela 3), tym samym będzie ono realizowane dłużej.



Rys. 7. Harmonogram projektu w punkcie kontrolnym $k = 1$ przed aktualizacją dostępności zasobów

Po uwzględnieniu dostępności zasobów i priorytetów otrzymujemy harmonogram (rys. 8), zgodnie z którym zostanie rozpoczęty projekt. Brakujące zapotrzebowanie na zasoby dla zadania (8, 9) zostało przesunięte na najbliższy możliwy termin w projekcie. Tym samym realizacja tego zadania wydłużyła się o dwa dni.



Rys. 8. Harmonogram projektu w punkcie kontrolnym $k = 1$ po aktualizacji dostępności zasobów

Zaktualizowany czas trwania zadania (i, j) obliczony w punkcie kontrolnym k projektu na podstawie wybranej strategii oraz informacji o dostępnych zasobach definiujemy przez KR_{ij}^k . Zaktualizowane czasy trwania zadań przedstawia tabela 5.

Tabela 5

Zaktualizowane o dostępność zasobów czasy trwania zadań KR_{ij}^1
w punkcie kontrolnym projektu $k = 1$

Zadanie (i, j)	K_{ij}^1	KR_{ij}^1
(7, 8)	14	14
(8, 9)	13	15
(1, 3)	3	3
(2, 3)	4	4
(3, 4)	4	4
(4, 5)	9	9
(5, 6)	11	11

Zgodnie z przyjętą strategią, realizacja zadania (8, 9) została zaplanowana na 15 dni, a realizacja zadania (4, 5) na 9 dni. Opóźnienie bądź przyspieszenie realizacji zadania (4, 5) będzie wpływało na zmianę czasu realizacji zadania (8, 9). Warto zwrócić uwagę, iż strategia optymistyczna dla zadania (4, 5) przewiduje 8 dni na realizację (również z uwzględnieniem zasobów). Załóżmy, że kierownik projektu ma informację o możliwym przyspieszeniu prac w tym zadaniu. W tym momencie mógłby on skorzystać ze strategii optymistycznej i zarezerwować zasoby tylko na 8 dni dla tego zadania. Dla potrzeb przykładu pozostaniemy jednak przy strategii uśrednionej dla wszystkich zadań.

5.2. Realizacja

W trakcie realizacji projektu (po 16 dniu) w punkcie kontrolnym $k = 2$ została dokonana aktualizacja czasów trwania zadań (patrz 4.1) na podstawie informacji o planowanym pp_{ij}^2 i zrealizowanym pr_{ij}^2 procencie zadań. Załóżmy, że przypuszczenie kierownika projektu z punktu kontrolnego $k = 1$ o szybszej realizacji zadania (4, 5) się sprawdziło (tabela 6).

uwzględnienie strategii innej niż uśredniona, jak np. optymistyczna czy pesymistyczna, może przynieść wymierne korzyści przy zarządzaniu zasobami w projekcie.

Podsumowanie

W pracy zaproponowano metodę harmonogramowania (proaktywnego i reaktywnego) projektów, w których czasy zadań nie są znane dokładnie, ponadto występują ograniczenia na dostępność zasobów. Czasy realizacji zadań modeluje się w postaci liczb rozmytych – zarówno na etapie planowania projektu, jak i w trakcie jego realizacji (dotyczy to oczywiście zadań jeszcze niezakończonych). Liczba rozmyta reprezentuje wszystkie możliwe czasy trwania zadania, wraz z ich stopniami możliwości. Decydent, na podstawie aktualnej wiedzy oraz swojego stosunku do ryzyka, musi dokonać wyboru deterministycznego oszacowania, przy czym wybór ten może być przez niego modyfikowany w trakcie realizacji projektu. Przydział zasobów jest dokonywany (i modyfikowany) na podstawie wyborów dokonywanych przez decydenta oraz zapotrzebowania i dostępności zasobów, a także priorytetów zadań. Podczas przydziału zasobów planowany czas realizacji zadania może się wydłużyć z powodu braku zasobów. W poszczególnych punktach kontrolnych decydent otrzymuje aktualne oszacowanie czasu realizacji projektu, uwzględniające aktualnie dostępne informacje oraz preferencje decydenta.

Dalszym kierunkiem badań może być opracowanie algorytmu automatyzacji aktualizacji wybranej strategii. Polegałby on na dopasowywaniu strategii wyboru deterministycznego czasu trwania zadania do zebranych dotychczas informacji dotyczących jego realizacji. Innymi słowy, mechanizm ten wspomagałby decydenta przy wyborze nowej deterministycznej wartości czasu trwania zadania.

Literatura

1. Chanas S., Zieliński P. (2001). Critical Path Analysis in the Network with Fuzzy Activity Times. *Fuzzy Sets and Systems*, 122, s. 195-204.
2. Chen S-J., Hwang CH-L. (1997). *Fuzzy Multiple Attribute Decision Making*. Springer Verlag, Berlin-Heidelberg.
3. Chen S-P. (2007). Analysis of Critical Paths in a Project Network with Fuzzy Activity Times. *European Journal of Operational Research*, 183, s. 442-429.

4. Duc Long L., Ohsato A. (2008). Fuzzy Critical Chain Method for Project Scheduling under Resource Constraints and Uncertainty. *International Journal of Project Management*, 26, s. 688-698.
5. Janczura M., Kuchta D. (2010). Reactive Scheduling with Fuzzy Activity Times. 31th International Conference ISAT2010 (Information System, Architecture, and Technology). Szklarska Poręba, 19-21 wrzesień.
6. Klimek M. (2010). Predyktywne harmonogramowanie projektów informatycznych. Konferencja Informatyki Stosowanej. Chełm 28.05.
7. Kuchta D. (2001). Miękka matematyka w zarządzaniu. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
8. Lambrechts O., Demeulemeester E., Herroelen W. (2007). Proactive and Reactive Strategies for Resource-constrained Project Scheduling with Uncertain Resource Availabilities. Springer.
9. Negnevitsky. (2002). Artificial Intelligence: A Guide to Intelligent Systems. Pearson Education.
10. Wang J. (2002). A Fuzzy Project Scheduling Approach to Minimize Schedule Risk for Product Development. *Fuzzy Sets and Systems*, 127, s. 99-116.

REACTIVE SCHEDULING AND RISK IN THE PROJECT

Summary

A project realization is usually dependent on many unexpected events. These risks have a significant impact on the project final total duration time.

This paper describes an algorithm based on a new method of project rescheduling [5]. To cope with an uncertainty of activity times the fuzzy numbers are used. The method provides a reactive scheduling strategy, based on updating the project schedule while the project realization is in progress. In this paper we improve the method by considering the resource constraint problem. Our mechanism is prepared to react to risks occurring in the project scheduling and running phase. We show how dialog with a user allows for making necessary decisions in response to unexpected problems. In order to illustrate the proposed approach, a simulation of a sample project realization is used.

Dorota Kuchta

Politechnika Wrocławska

REAKTYWNE HARMONOGRAMOWANIE PROJEKTÓW W PRZYPADKU ROZMYTYCH CZASÓW TRWANIA ZADAŃ I ROZMYTYCH LIMITÓW DOSTĘPNOŚCI ZASOBÓW

Wprowadzenie

Zarządzanie projektami jest w coraz większym stopniu obarczone niepewnością. Dlatego zarówno planowanie projektów powinno być przeprowadzane w sposób uwzględniający ten fakt, jak również kontrola realizacji projektów powinna być bardziej systematyczna i rzetelna. Konieczny jest wczesny system ostrzegania, który wskaże na potencjalne skutki problemów, zanim te skutki faktycznie nastąpią.

Niniejsze opracowanie poświęcone jest kontroli realizacji projektów, w których niepewne są zarówno czasy trwania zadań, jak i dostępność zasobów w poszczególnych okresach. Niepewne wartości modelowane są za pomocą liczb rozmytych. W poszczególnych punktach kontrolnych przeprowadzana jest częściowo zautomatyzowana aktualizacja planów, uwzględniająca uaktualniony poziom ryzyka oraz aktualne informacje, dostarczone przez decydenta. Zaproponowany model będzie spełniał rolę „systemu ostrzegania”, gdyż w każdym punkcie kontrolnym będzie wyznaczany przewidywany termin zakończenia projektu, uwzględniający aktualną sytuację, aktualnie dostępne informacje oraz akceptowany poziom ryzyka.

Najpierw przedstawimy podstawowe informacje dotyczące liczb rozmytych, stosowanych w proponowanym modelu, następnie zaprezentujemy zarys procesu planowania projektu. Potem zaproponujemy metodę aktualizacji planu projektu w odpowiedzi na zmiany, jakie zachodzą w warunkach realizacji zadania, czyli reaktywnego harmonogramowania projektów. Zaproponowana metoda zostanie zilustrowana na przykładzie.

1. Podstawowe informacje o zbiorach rozmytych

W niniejszym opracowaniu do modelowania niepewności użyjemy liczb rozmytych. Będą to liczby rozmyte specjalnego typu, tzw. liczby jednostronne [5], które pozwalają modelować ryzyko zwiększenia wartości, która powinna nie być zbyt duża, lub zmniejszenia wartości, która nie powinna być zbyt mała. Oba typy liczb rozmytych będą reprezentowane przed pary liczb, zapisywane będą jednak w postaci trójek liczb, aby ich interpretacja była jednoznaczna. Prawostronne liczby rozmyte będą reprezentowane przez trójkę liczb (a, a, b) , gdzie $b \geq a$, a różnica $b - a$ będzie przedstawiać stopień ryzyka, że wielkość reprezentowana przez liczbę rozmytą (a, a, b) , która w zasadzie powinna być równa a , będzie jednak większa (co jest uważane za niekorzystne). Stopień możliwości, że modelowana wielkość będzie równa a , uznawany jest za równy 1, a dla wartości z przedziału $(a, b]$ maleje on liniowo do 1 do 0. Analogicznie definiowane są lewostronne liczby rozmyte – są to trójki (a, b, b) , gdzie $b \geq a$ i stopień możliwości bycia faktyczną wartością ocenianej wielkości (która w zasadzie powinna być równa b i nie powinna być zbyt mała) rośnie w przedziale $[a, b]$ od 0 do 1. W obu przypadku poza przedziałem $[a, b]$ znajdują się wartości, których wystąpienie w przypadku modelowanej wielkości uznane jest za niemożliwe.

2. Proaktywne i reaktywne harmonogramowanie projektów

Rozpatrujemy H momentów czasowych, $t = t_0, t_1, \dots, t_H$, gdzie t_0 jest momentem planowania projektu, a t_H jest maksymalnym horyzontem czasowym – realizacja projektu absolutnie nie może przeciągnąć się poza ten moment. Zakładamy, że poszczególne zadania mogą się zaczynać i kończyć tylko w momentach $t = t_0, t_1, \dots, t_H$. W momencie t_0 przeprowadzamy planowanie projektu, a w momentach $t_1, \dots, t_{H-1}$ przeprowadzamy kontrolę realizacji i dokonujemy ewentualnej korekty planu.

Przyjmujemy, że w momencie t_0 wiemy, jakie zadania mają być wykonane w projekcie $(A_i, i=1, \dots, N)$, i zakładamy, że ten zbiór się nie zmieni – do końca projektu będzie się składał z tych samych elementów. Zakładamy również, że relacje poprzedzania między zadaniami są również znane na początku i ustalone do zakończenia projektu.

Zakładamy, że każde zadanie ma pięć atrybutów. Dwa z nich (są to wartości planowane) mogą się zmieniać w czasie i w pewnym momencie (kiedy znane są wartości rzeczywiste) stają się nieistotne, dwa inne (wartości rzeczywiste) są albo nieznane, albo ostatecznie ustalone, jeden (zapotrzebowanie na zasoby) będzie w tym artykule uznawany za stały, ale w ogólnym przypadku może się zmieniać w czasie. Oto oznaczenia tych atrybutów (dla $i=1, \dots, N$, $j=1, \dots, H$):

1. $PS_i(j)$ – początek zadania A_i planowany w momencie t_j . Jeśli w momencie t_j zadanie już zostało rozpoczęte, $PS_i(j)$ jest nieistotne, ponieważ znany jest rzeczywisty początek zadania.

2. $PD_i(j)$ – czas trwania zadania A_i planowany w momencie t_j . Jeśli w momencie t_j zadanie już zostało zakończone, $PD_i(j)$ jest nieistotne, ponieważ znany jest rzeczywisty początek zadania. Tutaj zakładamy, iż po rzeczywistym rozpoczęciu zadania już jest ustalone i nie będzie ulegało zmianie, ale z tego założenia można łatwo zrezygnować.

3. $AS_i(j)$ – rzeczywisty moment rozpoczęcia zadania A_i określony w momencie t_j . Jeśli w momencie t_j atrybut ten nie jest określony, oznacza to, że zadanie nie zostało rozpoczęte. Jeśli zadanie w momencie t_j już zostało rozpoczęte, $AS_i(j)$ jest ustalone dla wszystkich $t_l \geq t_j$, $l=1, \dots, H$. Oczywiście mamy wtedy $t_j \geq AS_i(j)$.

4. $AD_i(j)$ – rzeczywisty czas realizacji zadania A_i określony w momencie t_j . Jeśli zadanie w momencie t_j już zostało zakończone, $AD_i(j)$ jest ustalone dla wszystkich $t_l \geq t_j$, $l=1, \dots, H$, w przeciwnym przypadku wartość ta nie jest określona.

5. RR_i – liczba jednostek zasobu wymagana przez zadanie A_i . W tym artykule zakładamy, że wartość ta nie zmieni się, stąd nie rozróżniamy wartości planowanej i rzeczywistej, w ogólnym przypadku atrybut ten też może ulegać zmianie w stosunku do planów. W ogólnym przypadku możemy też mieć do czynienia z wieloma typami zasobów, tutaj zakładamy, że zasoby są nierozróżnialne. W ogólnym przypadku zapotrzebowanie zadania na zasoby też może się zmieniać w poszczególnych okresach realizacji zadania, tutaj zakładamy, iż jest ono stałe.

W modelu należy uwzględnić limit dostępności zasobu w projekcie. Zakładamy, iż dostępność ta może być różna w różnych okresach $[t_l, t_{l+1})$, $l=0, \dots, H-1$. Dodatkowo przyjmujemy, że wiedza na temat limitu dostępności

zasobów w poszczególnych okresach czasu z przyszłości może się zmieniać, w miarę zdobywania kolejnych informacji. Liczba jednostek zasobu dostępna w okresie $[t_l, t_{l+1})$ oceniana w momencie t_j ($t_l > t_j$) będzie oznaczana jako $R(l, j)$.

Proaktywne harmonogramowanie projektu odnosi się do momentu t_0 . Zazwyczaj (choć możliwe są inne funkcje celu) polega ono na określeniu wartości $PD_i(0)$, a następnie wartości $PS_i(0)$ w taki sposób, by planowany czas wykonania projektu był jak najkrótszy, przy czym zachowane muszą być wszystkie relacje poprzedzania między zadaniami i we wszystkich okresach $[t_l, t_{l+1})$, $l=0, \dots, H-1$ suma jednostek zasobu wymaganych przez takie zadania A_i , że $PS_i(0) \leq t_l$ i $PS_i(0) + PD_i(0) \geq t_{l+1}$ nie może przekraczać $R(l, 0)$. Problemowi harmonogramowania na etapie planowania poświęcona jest niezwykle bogata literatura [4]. Ponieważ problem jest obliczeniowo złożony, konstruuje się wiele algorytmów heurystycznych. Zazwyczaj zakłada się, że wszystkie wartości $PD_i(0)$ są znane. Jednak, ponieważ estymacja tych parametrów to problem sam w sobie, również jemu poświęca się uwagę, biorąc po uwagę aspekty psychologiczne oraz uwzględniając ryzyko błędu poprzez stosowanie w harmonogramie buforów [2; 6].

Podczas realizacji projektu harmonogram wygenerowany w momencie t_0 zazwyczaj okazuje się w znacznym stopniu inny niż harmonogram faktycznie realizowany. Wynika to z błędów lub niewiedzy na etapie planowania. Podczas realizacji projektu pojawiają się nowe informacje, mają miejsce nieprzewidziane zdarzenia, efektywność pracy zespołu okazuje się być inna niż zakładano itp. Dlatego konieczne jest harmonogramowanie reaktywne – czyli ponowne harmonogramowanie niewykonanej jeszcze „reszty” projektu w momentach $t=t_1, \dots, t_H$. Literatura poświęcona temu problemowi jest znacznie mniej liczna [7]. Zakłada się, że w momentach $t=t_1, \dots, t_H$ można po prostu zastosować te same algorytmy, co w momencie planowania projektu, tyle że przedmiotem działania tych algorytmów będzie projekt „okrojony” o już wykonaną część. Trzeba jednak wziąć pod uwagę fakt, że podczas realizacji projektu pojawiają się nowe informacje, warunki realizacji projektu mogą się zmienić i to wszystko musi być wzięte pod uwagę przy aktualizacji harmonogramu projektu. Ponadto, podczas realizacji projektu jest zazwyczaj znacznie mniej czasu niż w momencie t_0 , dlatego metody stosowane przy reaktywnym harmonogramowaniu muszą być w dużym stopniu zautomatyzowane, dopuszczając jednak w każdym momencie interwencję decydenta (czyli wprowadzenie nowych informacji w trybie interaktywnym). Dlatego potrzebne są metody przewidziane specjalnie

dla harmonogramowania reaktywnego. Jak wspomniano, takie metody są zaproponowane w literaturze [7]. Dotyczą one jednak przypadku, kiedy wartości $PD_i(j)$ i $R(l,j)$, $j=0,...,H$, $l=0,...,H-1$ są deterministyczne.

Jednak wiadomo powszechnie, iż zazwyczaj wartości $PD_i(j)$, $j=0,...,H-1$ nie mogą być podane od razu w postaci deterministycznej. W większości przypadków niezbędne jest uwzględnienie pewnego stopnia ryzyka. To samo dotyczy, często w jeszcze większym stopniu, wartości $R(l,j)$, $j=0,...,H$, $l=0,...,H-1$. Dostępność zasobów ludzkich zmienia się z powodu chorób, wyjazdów, niezaplanowanych zadań, w przypadku zasobów materialnych w grę wchodzi awarie, w każdym przypadku problemem jest dzielenie zasobów przez różne projekty w organizacji i ich wzajemne „wyrywanie” sobie. W celu modelowania niepewności stosuje się podejście probabilistyczne lub rozmyte. W tym opracowaniu użyjemy tego ostatniego, ponieważ zdaniem wielu autorów jest ono prostsze do zastosowania w praktyce.

3. Zastosowanie liczb rozmytych do harmonogramowania proaktywnego i reaktywnego

Podejście rozmyte jest często stosowane w literaturze do harmonogramowania projektów [1]. Nie zostało ono wszakże prawie wcale zastosowane do harmonogramowania reaktywnego. Autorce znana jest jedna tylko praca [3], w której w podejściu reaktywnym stosowane jest podejście rozmyte. Wydaje się, iż powszechnie przyjmuje się założenie, że podczas realizacji projektu, w momentach $t=t_1,...,t_H$, znane są deterministyczne wartości odpowiednich parametrów (podejście probabilistyczne również nie jest stosowane w harmonogramowaniu reaktywnym). Jednak, zwłaszcza w projektach o stosunkowo długim czasie trwania, w momentach t_j , istotnie mniejszych od planowanego końca projektu, podawane wartości deterministyczne prawie zawsze obarczone będą błędem – podobnie jak to było na etapie planowania, w momencie t_0 . Dlatego celem niniejszego opracowania jest zaproponowanie metody harmonogramowania proaktywnego i reaktywnego w sytuacji, kiedy wartości $PD_i(j)$ and $R(l,j)$, $j=0,...,H-1$, $l=0,...,H-1$ mogą być liczbami rozmytymi.

Różnica między propozycją zawartą w niniejszym opracowaniu a metodą zaproponowaną przez Janczurę i Kuchtę [3] polega na dopuszczeniu rozmytości również w przypadku wartości $R(l,j)$, $j=0,...,H$, $l=0,...,H-1$ (we wspomnianej pracy limity na dostępność zasobów w ogóle nie są uwzględniane), a przede wszystkim na innym sposobie wyznaczania wartości $PD_i(j)$, $j=0,...,H-1$.

Janczura i Kuchta [3] zakładają, że wartości $PD_i(j)$, $j=1, \dots, H-1$ zależą od przebiegu realizacji i -tego zadania do momentu t_j . Jeśli w momencie t_j nie ma informacji o przebiegu realizacji i -tego zadania, czyli jeśli zadanie to nie zostało rozpoczęte, mamy równość $PD_i(j) = PD_i(0)$. Tutaj rozpatrujemy projekty o znacznie dłuższym czasie trwania i modyfikujemy, w razie potrzeby, informację o planowanym czasie realizacji danego zadania systematycznie, często na długo zanim zadanie to zostanie rozpoczęte – aby w każdym momencie kontrolnym mieć możliwie rzetelną prognozę przewidywanego terminu zakończenia całego projektu. Zakładamy, że im bliżej aktualnie przewidywanego terminu rozpoczęcia danego zadania, tym dokładniej możemy określić jego przewidywany czas realizacji. Co więcej, przyjmujemy, że czas realizacji poszczególnych zadań jest na tyle krótki w stosunku do czasu realizacji całego projektu, iż w momencie rozpoczęcia realizacji danego zadania znamy już dokładnie jego czas realizacji. To założenie łatwo można anulować, łącząc proponowane tu podejście z podejściem zaproponowanym przez Janczurę i Kuchtę [3], gdzie właśnie rzeczywista realizacja zadania decyduje o uaktualnieniach jego przewidywanego czasu realizacji. Wartości $R(l,j)$, nieuwzględniane w pracy Janczury i Kuchty, są tutaj uaktualniane również przy założeniu, że im mniejsza różnica $l-j$, tym mniejsza niepewność jest związana z $R(l,j)$.

3.1. Proponowane zastosowanie liczb rozmytych do harmonogramowania proaktywnego

Niniejszy podrozdział nie zawiera żadnych nowych elementów, jednak, żeby mówić o harmonogramowaniu reaktywnym, trzeba mieć początkowy planowany harmonogram projektu. Tutaj skonstruujemy harmonogram bazujący na rozmytych oszacowaniach czasów trwania zadań i limitów dostępności zasobów.

Harmonogramowanie proaktywne jest przeprowadzane w momencie t_0 . Niepełną informację o przewidywanych czasach trwania zadań i o limitach dostępności zasobów wyrazimy w postaci jednostronnych liczb rozmytych typu (b,b,a) lub (a,b,b) .

Zakładamy, że wartości $PD_i(0)$ są prawostronnymi liczbami rozmytymi, oznaczanymi jako $\tilde{PD}_i(0) = (\underline{pd}_i(0), \underline{pd}_i(0), \overline{pd}_i(0))$, gdzie różnica $\overline{pd}_i(0) - \underline{pd}_i(0)$ wyraża niepewność w stosunku do wartości $\underline{pd}_i(0)$, uważanej w momencie t_0 za najbardziej możliwy czas trwania i -tego zadania. Podobnie, wartości $R(l,0)$ będą lewostronnymi liczbami rozmytymi $\tilde{R}_l(1,0)$, $l=0, \dots, H-1$,

mającymi postać $(r_l(0), \bar{r}_l(0), \bar{r}_l(0))$, gdzie $\bar{r}_l(0)$ oznacza najbardziej możliwą wartość limitu zasobów w okresie $[t_l, t_{l+1})$, $l=1, \dots, H-1$, a różnica $\bar{r}_l(0) - r_l(0)$ wyraża stopień ryzyka (niepewności), że w rzeczywistości limit ten będzie mniejszy.

Kiedy znane są oszacowania $PD_i(0)$ i $\tilde{R}_i(l,0)$ możemy zastosować jeden z algorytmów znanych w literaturze, wyznaczający planowany harmonogram w przypadku rozmytych czasów trwania zadań i rozmytych dostępności zasobów. Niektóre z tych algorytmów wyznaczają harmonogram o rozmytych planowanych terminach rozpoczęcia zadań, niektóre wyznaczają harmonogram deterministyczny. My preferujemy to drugie podejście, z uwagi na fakt, że zespół projektowy zazwyczaj oczekuje konkretnej informacji co do planowanego sposobu działania. Rozmytość powinna być znana i szacowana przez osoby zajmujące się planowaniem i kontrolą realizacji projektu, ale niekoniecznie musi być przekazywana zespołowi w ramach informacji o harmonogramie pracy. Dlatego liczby rozmyte $PD_i(0)$ i $\tilde{R}_i(l,0)$ należy poddać „defuzyfikacji”, aby otrzymać ich deterministyczne odpowiedniki $PD_i(0)$ i $R_i(l,0)$. Można to zrobić na wiele sposobów, niżej wykorzystujemy jeden z nich.

W celu ilustracji proponowanego podejścia, na etapie harmonogramowania proaktywnego zastosujemy następujący prosty algorytm, wyznaczający deterministyczny harmonogram w momencie t_0 . Wartości $PS_i(0)$ dla każdego $i=1, \dots, N$ będą zatem liczbami deterministycznymi, wyznaczonymi na podstawie rozmytych informacji o projekcie i preferencji decydena.

Algorytm I

Krok 0. Wybierz akceptowany poziom ryzyka λ z przedziału $[0,1]$ (0 – ryzyko nieakceptowane w żadnym stopniu, 1 – akceptowane pełne ryzyko, zazwyczaj λ będzie bliskie 0).

Krok 1. Przyjmij wartość $PD_i(0) = r_l(0) + (1 - \lambda)(\bar{r}_l(0) - r_l(0))$ jako oszacowanie czasu trwania zadania dla $i=1, \dots, N$ oraz $R_l(0) = \bar{r}_l(0) - (1 - \lambda)(\bar{r}_l(0) - \bar{r}_l(0))$ jako oszacowanie limitu dostępności zasobów w poszczególnych okresach $[t_l, t_{l+1})$, $l=1, \dots, H-1$.

Krok 2. Zastosuj dowolny algorytm minimalizujący czas realizacji projektów w warunkach ograniczonej dostępności zasobów w przypadku deterministycznym [7], do wyznaczenia $PS_i(0)$ dla każdego $i=1, \dots, N$. Danymi wejściowymi dla algorytmu są wartości $PD_i(0)$ i $R_l(0)$, $i=1, \dots, N$ oraz relacje poprzedzania między zadaniami.

Algorytm I jest algorytmem „ostrożnym”, odpowiadającym postawie pesymisty. Przyjmuje się niemal największe możliwe czasy trwania zadań i niemal najmniejsze możliwe limity dostępności zasobów. „Niemal” – ponieważ dopuszczany jest pewien poziom ryzyka. Jak już wspomniano, jest to tylko jeden z możliwych algorytmów i można by w tym miejscu zastosować każdy inny algorytm, który wyznacza deterministyczny harmonogram w przypadku rozmytych czasów trwania zadań i rozmytych limitów dostępności zasobów.

3.2. Proponowane zastosowanie liczb rozmytych w harmonogramowaniu reaktywnym

W momencie t_0 , na etapie planowania, jest, a przynajmniej powinno być, dostatecznie dużo czasu na estymację wartości $P\tilde{D}_i(0)$ and $\tilde{R}_i(l,0)$. Jednak doświadczenie pokazuje, że w trakcie realizacji projektu czasu na reestymację odpowiednich wartości jest znacznie mniej. Dlatego istnieje potrzeba częściowo automatycznego generowania kolejnych wartości $P\tilde{D}_i(j)$, $\tilde{R}_i(l,j)$, $j=1,...,H-1$, $l=2,...,H-1$, $l>j$, przy czym oczywiście decydent miałby w każdym punkcie kontrolnym t , $t=t_j$, $j=1,...,H-1$, możliwość rezygnacji z automatycznych uaktualnień i mógłby wprowadzić własne wartości, wynikające z jego aktualnej wiedzy. Wartości $P\tilde{D}_i(j)$, $\tilde{R}_i(l,0)$ będą również mogły zostać zdefuzyfikowane, czy to dlatego, że będzie dostępna odpowiednia wartość dokładna, czy to na potrzeby algorytmu aktualizacji harmonogramu (algorytm III). Będą wówczas oznaczane odpowiednio $PD_i(j)$, $R_i(l,0)$

Zakładamy w każdym momencie t_j , $j=1,...,H-1$, że im bliższy jest planowany moment $PS_i(j)$ rozpoczęcia i -tego zadania, tym lepiej umiemy oszacować czas trwania tego zadania, a w momencie jego rzeczywistego rozpoczęcia czas ten umiemy oszacować dokładnie. Jak wspomniano wyżej, założenie to jest uzasadnione w przypadku projektów długoterminowych, przy których stosunek czasu trwania poszczególnych zadań do czasu trwania całego projektu jest mały. Analogiczne założenie przyjmujemy w stosunku do oszacowań limitów na dostępność zasobów w poszczególnych okresach: im bliżej danego okresu, tym lepiej ją znamy.

W każdym momencie t_j , $j=1, \dots, H-1$ proponujemy zatem zastosować następujący algorytm:

Algorytm II

Krok 0. $i=1$.

Krok 1. Jeśli zadanie A_i do momentu t_j włącznie zostało rozpoczęte, ale nie zostało zakończone, przyjmij za $PS_i(j)$ i $AS_i(j)$ rzeczywisty moment rozpoczęcia zadania, a za deterministyczne $PD_i(j)$ wartość podaną przez eksperta lub $\underline{pd}_i(j-1)$.

Krok 2. Jeśli zadanie A_i do momentu t_j włącznie zostało zakończone, przyjmij za $PS_i(j)$ i $PD_i(j)$ oraz $AS_i(j)$ i $AD_i(j)$ odpowiednie wartości rzeczywiste.

Krok 3. Jeśli do momentu t_j włącznie zadanie A_i nie zostało rozpoczęte, oszacuj jego planowany czas trwania za pomocą liczby rozmytej $\tilde{PD}_i(j) = (\underline{pd}_i(j), \underline{pd}_i(j), \overline{pd}_i(j))$, gdzie $\underline{pd}_i(j)$ i $\overline{pd}_i(j)$ są albo podane przez eksperta, albo wyznaczone z następujących wzorów

$$\underline{pd}_i(j) = \underline{pd}_i(j-1) \quad (1)$$

$$\overline{pd}_i(j) = \underline{pd}_i(j) + \frac{PS(j-1) - t_i}{PS(j-1)} (\overline{pd}_i(j-1) - \underline{pd}_i(j-1)) \quad (2)$$

Krok 4. Oceń $\tilde{R}(l, j)$ dla $l > j$ jako liczbę rozmytą $(\underline{r}_l(0), \bar{r}_l(0), \bar{r}_l(0))$, gdzie $\underline{r}_l(0)$ i $\bar{r}_l(0)$ mogą być podane przez eksperta albo wyznaczone z poniższych formuł

$$\bar{r}_l(j) = \bar{r}_l(j-1) \quad (3)$$

$$\underline{r}_l(j) = \bar{r}_l(j) - \frac{l-j}{l} (\bar{r}_l(j-1) - \underline{r}_l(j-1)) \quad (4)$$

Krok 5. Jeśli $i < N$, podstaw $i=i+1$ i idź do Kroku 1. W przeciwnym przypadku STOP.

Wzory (1) i (2) wynikają z następującego założenia: im bliżej planowanego momentu rozpoczęcia danego zadania, tym mniejsza jest niepewność z nim związana. Jeśli ekspert nie poda żadnej nowej informacji, przyjmuje się, że oszacowanie optymistyczne czasu trwania zadania (wzór (1)) pozostaje bez zmian, natomiast niepewność związana z tym oszacowaniem zmniejsza się proporcjonalnie do czasu, jaki upłynął od ostatniego dnia kontroli. Na analogicznych założeniach opierają się oszacowania dotyczące zasobów – wzory (3) i (4).

Kiedy dysponujemy rozmytymi oszacowaniami czasów trwania zadań i limitów dostępności zasobów dla niezrealizowanej jeszcze części projektu, możemy wyznaczyć zaktualizowany deterministyczny harmonogram dalszej realizacji projektu, stosując np. modyfikację algorytmu I. Modyfikację tę zastosujemy w poszczególnych momentach t_j , $j=1, \dots, H-1$ do zaktualizowanego projektu, czyli do projektu charakteryzowanego przez wartości $PS_i(j)$ lub $\tilde{PS}_i(j)$, $PD_i(j)$ lub $\tilde{PD}_i(j)$ i $\tilde{R}(l,j)$ dla $l > j$:

Algorytm III

Krok 0. Wybierz akceptowany poziom ryzyka λ z przedziału $[0,1]$.

Krok 1:

- dla wszystkich $i=1, \dots, N$ takich, że A_i do momentu t_j nie zostało rozpoczęte, przyjmij $PD_i(j) = \underline{r}_i(0) + (1-\lambda)(\bar{r}_i(0) - \underline{r}_i(0))$,
- przyjmij $R_l(j) = \bar{r}_l(0) - (1-\lambda)(\bar{r}_l(0) - \bar{r}_l(0))$ dla każdego okresu $[t_l, t_{l+1})$, $l=t+1, \dots, H-1$.

Krok 2. Zastosuj dowolny algorytm minimalizujący czas realizacji projektu w przypadku ograniczeń na dostępność zasobów (dla deterministycznych parametrów), w celu wyznaczenia wartości $PS_i(j)$ dla każdego $i=1, \dots, N$ takiego, że zadanie A_i nie zostało rozpoczęte – danymi wejściowymi dla tego algorytmu są wartości rzeczywiste dla już zakończonych zadań, wartości $PD_i(j)$ dla zadań jeszcze nie zakończonych oraz wartości $R_l(j)$ dla każdego okresu $[t_l, t_{l+1})$, $l=t+1, \dots, H-1$. Na wyjściu otrzymamy deterministyczne wartości $PS_i(j)$ dla każdego $i=1, \dots, N$ takiego, że zadanie A_i nie zostało rozpoczęte.

Złożoność obliczeniowa zaproponowanego rozwiązania jest taka, jak złożoność algorytmu wybranego w Kroku 2 – wśród dostępnych algorytmów są heurystyki o małej złożoności, ale także algorytmy dokładne o znacznej złożoności.

4. Przykład

Rozpatrzmy bardzo prosty przykład, ilustrujący ideę niniejszego artykułu. Mamy H (gdzie H jest odpowiednio duże) momentów czasowych, rozdzielonych równomiernie tak „gęsto”, na ile to jest uzasadnione i wymagane w danym przypadku, np. co $1/8$ okresu (wówczas mielibyśmy momenty czasowe $0, 0,125, 0,25, \dots$). Zadania mogą zaczynać się i kończyć tylko w tych momentach – nawet jeśli zadanie skończy się wcześniej, to jego zakończenie będzie uznane dopiero w najbliższym z wyróżnionych momentów czasowych. Niech projekt składa się z dwóch zadań, oznaczonych odpowiednio symbolami $A1$ i $A2$ i połączonych relacją (Koniec, Początek, -50%), gdzie $A2$ jest następnikiem $A1$. Relacja ta oznacza, że zadanie $A2$ może się rozpocząć nie wcześniej niż w momencie, kiedy zostanie zrealizowana połowa pracy przy zadaniu $A1$ – zatem częściowo zadania te mogą być wykonywane równoległe, jeśli wystarczy zasobów. Czasy trwania obu zadań w momencie planowania oceniamy w sposób następujący: $\tilde{PD}_1(0)=(2,5, 2,5, 2,5)$, $\tilde{PD}_2(0)=(2,2,3)$ – czas realizacji zadania $A2$ jest jeszcze obciążony niepewnością. Ponadto, liczba wymaganych zasobów w obu zadaniach wynosi po dwa, mamy zatem $RR_1=RR_2=2$. Jeśli chodzi o dostępność zasobów, to oceniamy ją dla $H=6$ i mamy: $R_0(0)=4$, $\tilde{R}_0(1)=(3,4,4)$, $\tilde{R}_0(2)=(2,5, 4, 4)$, $\tilde{R}_0(3)=\tilde{R}_0(4)=\tilde{R}_0(5)=(2,4,4)$ – zatem w zasadzie uważamy, że każdym momencie będą dostępne 4 jednostki zasobu, jednak informacja ta w momencie 0 nie jest pewna, jeśli chodzi o okresy następujące po momencie 1.

Za akceptowany poziom ryzyka w algorytmie I przyjmujemy $0,5$. Wówczas mamy: $PD_1(0)=2,5$, $PD_2(0)=2,5$ oraz $R_0(0)=4$, $R_0(1)=3,5$, $R_0(2)=3,75$, $R_0(3)=R_0(4)=R_0(5)=3$. W wyniku zastosowania algorytmu I powstanie harmonogram zawarty w tabeli 1.

Tabela 1

Plan realizacji przykładowego projektu, skonstruowany w momencie 0 przy akceptowanym ryzyku $0,5$

Okres	[0,1)	[1,2)	[2, 2,5)	[2,5, 3)	[3,4)	[4,5)
Realizowane zadania	$A1$	$A1$	$A1$	$A2$	$A2$	$A2$
Potrzebne zasoby	2	2	2	2	2	2
Dostępne zasoby	4	3,5	3,25	3,25	3	3

Wprawdzie, przy założeniu postępu prac przy zadaniu $A1$ proporcjonalnego do upływu czasu, rozpoczęcie zadania $A2$ mogłoby zostać zaplanowane na moment 1,25 (kiedy planowane jest ukończenie 50% zadania $A1$), ale w momencie 0 przyjmujemy, iż nie wiadomo, czy w tym momencie będą dostępne 4 jednostki zasobu. Dlatego zostaje zbudowany plan ostrożny, proaktywny, przy którym termin zakończenia jest prognozowany na poziomie ryzyka 0,5 i wynosi 5 miesięcy.

Założmy, że w momencie 1 jest przeprowadzana kontrola realizacji. Okazało się, że zostało zrealizowane 50% zadania $A1$ i przystąpiono do realizacji zadania $A2$. Wówczas, zgodnie z algorytmem II, otrzymamy następujące oceny limitów dostępności zasobów: $R_1(1)=4$, $\tilde{R}_1(2)=(3,5, 4, 4)$, $\tilde{R}_1(3)=(3, 4, 4)$, $\tilde{R}_1(4)=(2,5, 4,4)$, $\tilde{R}_1(5)=(2,4, 4, 4)$ oraz, zgodnie z opinią eksperta, deterministyczne czasy trwania obu zadań: $PD_1(1)=PD_2(1)=2$. Algorytm III, na poziomie ryzyka 0,5, pozwoli wyznaczyć skorygowany plan projektu – tabela 2.

Tabela 2

Reaktywny plan realizacji przykładowego projektu, skonstruowany w momencie 1 przy akceptowanym ryzyku 0,5

Okres	[1,2)	[2, 3)
Zadania realizowane	$A1, A2$	$A2$
Potrzebne zasoby	4	2
Dostępne zasoby	4	3,5

Zatem planowany czas realizacji projektu został skrócony, dzięki temu, że część projektu została zrealizowana, minęła jedna jednostka czasu i wiadomo, iż w okresie [1,2) będą dostępne 4 zasoby, zatem oba zadania istotnie mogą być realizowane równolegle. Niemniej jednak na początku projektu, kiedy dostępność zasobów była obciążona niepewnością, podanie terminu realizacji projektu równego trzem jednostkom czasu było obciążone ryzykiem, zatem zastosowano wówczas harmonogramowanie proaktywne – ostrożne.

Podsumowanie

Zaproponowano metodę zarządzania ryzykiem przekroczenia terminu realizacji projektów. Musi ona być poddana weryfikacji w praktyce. Trzeba jednak zaznaczyć, że jest to jedna z niewielu metod reaktywnego zarządzania projektami, które w obliczu dzisiejszej zmienności warunków gospodarczych i społecznych będzie coraz ważniejsze dla sukcesu projektów. Wydaje się nawet, iż w przypadku wielu projektów, szczególnie narażonych na zmiany, jest ono nawet ważniejsze niż tak szeroko opisywane w literaturze naukowej i praktycznej planowanie proaktywne, czyli przeprowadzane przed rozpoczęciem projektu. Często nie bardzo wiadomo, co ma powstać, czego chce klient, a co umiemy zrobić. Takim projektem trzeba zarządzać reaktywnie, bo początkowe plany zapewne nie na wiele się zdadzą. Wydaje się zatem, iż niniejsze opracowanie wpisuje w szerszy nurt zarządzania projektami, który koniecznie musi być rozwijany.

Literatura

1. Chanas S., Zieliński P. (2001). Critical Path Analysis in the Network with Fuzzy Activity Times. *Fuzzy Sets and Systems*, 122, s. 195-204.
2. Goldratt E. (1997). *Critical Chain*. North River Press, Great Barrington.
3. Janczura M., Kuchta D. (2010). Reactive Scheduling with Fuzzy Activity Times. 31th International Conference ISAT2010 (Information System, Architecture, and Technology). Szklarska Poręba, 19-21 wrzesień.
4. Klimek M. (2010). Predyktywne harmonogramowanie projektów informatycznych. Konferencja Informatyki Stosowanej, Chełm 28.05.
5. Kuchta D. (2001). *Miękka matematyka w zarządzaniu*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
6. Kuchta D., Kobyłański P. (2007). A Note on the Paper by M. A. Al-Fawzan and M. Haouari about a Bi-objective Problem for Robust Resource-constrained Project Scheduling. *International Journal of Production Economics*, 107(2), s. 496-501.
7. Lambrechts O., Demeulemeester E., Herroelen W. (2007). Proactive and Reactive Strategies for Resource-constrained Project Scheduling with Uncertain Resource Availabilities. Springer.

REACTIVE SCHEDULING OF PROJECTS IN CASE OF FUZZY TASKS DURATION TIMES AND FUZZY RESOURCES AVAILABILITY LIMITS

Summary

Projects are carried out today in a dynamically changing environment. That is why information used during project planning may turn out to be useless during project realization. In the present paper it was assumed that the longer the time period between the planning moment and the task realization, the greater the uncertainty degree linked to the task and to the circumstances, in which the task will be realized. Hence an half interactive method is proposed of determining the initial project schedule and of updating it during project realization, taking into account the moment of schedule building/updating and the planned moment of task realization. It is assumed that the uncertainty concerns two aspects: the task duration and the resource availability.

Aleksandra Niedzielska

Politechnika Wrocławska

Dorota Kuchta

Politechnika Wrocławska

Wyższa Szkoła Oficerska Wojsk Lądowych im. gen. Tadeusza Kościuszki we Wrocławiu

PROBLEM ZAKUPÓW W DUŻYCH PROJEKTACH BADAWCZYCH

Wprowadzenie

Jednym z procesów zarządzania projektami jest proces zarządzania zakupami i dostawami. Proces ten, w zależności od charakteru projektu i zbioru potencjalnych dostawców, jest mniej lub bardziej istotny. Jednak przedmiotem niniejszego opracowania są projekty, gdzie proces ten ma ogromną wagę. Chodzi o wielkie projekty badawcze, w których zakupy opiewają na ogromne kwoty i są bardzo różnorodne, a potencjalni dostawcy mają swoje siedziby praktycznie na całym świecie, a już na pewno w całej Europie. Konkretnie zajmujemy się jednym z takich projektów. Analizujemy proces zarządzania zakupami w wybranym projekcie, by sprawdzić, jakie ma on wady i czy można zaproponować jego ulepszenie. Jest to niezwykle istotne: w grę wchodzi tutaj możliwość utraty wielkich kwot z powodu niewłaściwego wyrobu dostawcy, ale też sterowanie rozwojem tego czy innego państwa poprzez wybieranie lub niewybijanie firm z jego terytorium. Istnieje ryzyko podejrzeń, ale i faktycznych manipulacji oraz ryzyko niezdolności decydentów do właściwej oceny, z powodu odległości i zbyt dużej liczby oferentów.

Wybrany projekt będzie niżej opisany. Przeanalizowana zostanie rola polskich dostawców w tym projekcie, która po części jest konsekwencją stosowanego systemu wyboru dostawców (a po części skutkiem postawy polskich przedsiębiorców). Zostanie przeanalizowany system doboru dostawców, a następnie zostanie zaproponowana jego modyfikacja, wykorzystująca reguły rozmyte.

1. Projekt „Wielki Zderzacz Hadronów” – podstawowe informacje

Wielki Zderzacz Hadronów (Large Hadron Collider – LHC) jest to duży projekt naukowy (badawczy) wymagający budowy dużej infrastruktury badawczej. Projekt jest realizowany przez Europejską Organizację Badań Jądrowych (CERN). CERN jest międzynarodową organizacją naukową z siedzibą w Genewie, która została założona w 1954 roku z inicjatywy najbardziej rozwiniętych państw z Europy Zachodniej. Do Europejskiej Organizacji Badań Jądrowych (CERN) należy 20 krajów członkowskich oraz 8 państw obserwatorów. Polska stała się pełnoprawnym członkiem CERN 1 lipca 1991 roku. Dlatego też ma pełnoprawny dostęp do projektu naukowego LHC [3].

W CERN prowadzone są badania dotyczące struktury materii i własności cząstek elementarnych, w których wykorzystuje się najnowsze światowe osiągnięcia technologiczne i stymuluje się postęp naukowy w wielu dziedzinach gospodarki. CERN jest w stanie ze składek jej członków wybudować wielką infrastrukturę badawczą, która pozwala następnie realizować badania naukowe, których zakres i koszty przekraczają możliwości finansowe jednego kraju, nawet tego najbogatszego [3].

Dla krajów członkowskich obecność w tej organizacji niesie za sobą podwójne korzyści. Naukowcy mogą brać udział w badaniach na skalę światową, a przemysł może czerpać korzyści ekonomiczne uczestnicząc w pracach techniczno-produkcyjnych. Projekt LHC składa się z trzech części: przygotowania (projektu), budowy oraz eksploatacji. Całkowity budżet tego projektu do tej pory wyniósł około 4 mld euro. LHC jest to największy na świecie akcelerator cząstek elementarnych (hadronów). LHC jest zlokalizowany na granicy Szwajcarii i Francji w Europejskim Ośrodku Badań Jądrowych koło Genewy. Jest on najważniejszym i jedynym w skali światowej instrumentem badawczym współczesnej fizyki cząstek elementarnych. Fizycy wierzą, że projekt ten pomoże im odpowiedzieć na fundamentalne pytania w dziedzinie fizyki (np. dotyczące czarnej materii, bozonu Higgsa). Akcelerator cząstek LHC (jego zasadnicze elementy) jest umieszczony w torycznym tunelu o długości 27 km na głębokości od 50 m do 175 m pod ziemią. Projekt rozpoczął się w połowie lat 80. XX wieku i trwał do grudnia 1994 roku, kiedy to Rada CERN zatwierdziła program budowy LHC i projekt przeszedł w fazę budowy. Faza budowy LHC zakończyła się po 14 latach. 10 września 2008 roku nastąpiło jego uruchomienie. Jednak po 10 dniach akcelerator uległ awarii. W grudniu 2009 roku ponownie uruchomiono akcelerator. Projekt przeszedł ostatecznie w ostatnią fazę, czyli w fazę eksploatacji. Udział w projekcie biorą wszystkie kraje należące do europejskiej organizacji CERN [4; 8].

Jeśli chodzi o wybór dostawców, to bierze się pod uwagę oferowaną cenę oraz to, z jakiego kraju pochodzą dostawcy (czy jest to kraj zbilansowany czy niezbilansowany). Dąży się do tego, aby wartość dostaw do CERN była równa składce członkowskiej odprowadzanej przez poszczególne państwo – wtedy mówimy o zbilansowaniu danego kraju. Jeżeli zaś składka członkowska przewyższa wartość dostaw, mamy do czynienia z niezbilansowaniem danego kraju. Polska jest właśnie krajem niezbilansowanym [7].

2. Zakupy w projekcie „Wielki Zderzacz Hadronów”

W CERN zakupów dokonuje się tylko w krajach członkowskich. Przy wyłanianiu najlepszej oferty decyduje kryterium najniższej ceny, przy czym uwzględniana jest wcześniejsza preselekcja na podstawie kryteriów technicznych i ekonomicznych oferentów. W każdym pojedynczym zapytaniu o dostawę towarów lub usług określone są dokładne kryteria wyboru. W CERN przy realizacji projektu LHC wybór procedury wyboru wykonawcy jest zależny od spodziewanej ceny kontraktu liczonej w CHF:

- spodziewana cena kontraktu poniżej 10 000 CHF; zapytania o dostawę towarów lub usług są kierowane bezpośrednio przez dywizje techniczne CERN do konkretnych wykonawców, dlatego ważny jest tutaj bezpośredni kontakt z przedstawicielami CERN i zgłoszenie swojej oferty,
- spodziewana cena kontraktu powyżej 10 000 CHF i poniżej 200 000 CHF; obowiązuje procedura zapytania o cenę, tj. Price Enquires; pełną listę aktualnych zapytań o cenę można znaleźć na stronach internetowych albo bezpośrednio przez kontakt z odpowiednimi dywizjami w CERN, wykonawcy mają 4 tygodnie na odpowiedź od momentu ogłoszenia zapytania; ze względu na tak krótki okres, potencjalni wykonawcy powinni wcześniej nawiązać kontakt z przedstawicielami CERN,
- spodziewana cena kontraktu powyżej 200 000 CHF; obowiązuje procedura dwuetapowa, na kilka miesięcy wcześniej pojawia się w materiałach Komitetu Finansowego CERN zapowiedź wstępna.

Pierwszy etap to Przegląd rynku (Market Survey). Celem tego etapu jest wyłonienie firm, które zostaną następnie zaproszone do właściwego przetargu. Firmy, które odpowiedzą na Market Survey są oceniane pod względem formalnym oraz bywają także odwiedzane przez przedstawicieli CERN w celu przeprowadzenia audytu technologicznego. Firmy są dokładnie sprawdzane, czy są w stanie zrealizować kontrakt. Firmy, które pozytywnie przejdą ten etap, są następnie zaproszone do drugiego etapu, czyli Przetargu. Fakt, że dostawcy zazwyczaj są rozproszeni geograficznie, powoduje, że między pierwszym a drugim etapem może wystąpić kilkumiesięczna przerwa.

Drugi etap do Przetarg (Tender). Firmy, które pozytywnie przeszły pierwszy etap zostają zaproszone do złożenia swoich ofert w drugim etapie. Na tym etapie oferty są porównywane tylko na podstawie ceny. Po otwarciu ofert nie ma możliwości negocjacji cenowej. Oferta z najniższą ceną wygrywa. Istnieje odstępstwo od tej zasady, jeśli w przetargu biorą udział przedstawiciele z krajów niezbilansowanych. Jeśli najtańsza oferta pochodzi z kraju zbilansowanego, a oferta pochodząca z kraju niezbilansowanego jest droższa o nie więcej niż 20%, może zostać wybrany oferent pochodzący z kraju niezbilansowanego [7].

CERN dopuszcza wkłady rzeczowe „in-kind” do projektu LHC, ale tylko dla państw obserwatorów, które nie są członkami organizacji, lecz zadeklarowały wkład w LHC (np. USA, Rosja, Indie) [4].

3. Polscy dostawcy w projekcie „Wielki Zderzacz Hadronów”

Całkowity udział polskich dostaw dla LHC wyniósł do dnia dzisiejszego około 1% całkowitej wartości tej inwestycji, przy składce członkowskiej wynoszącej ponad 2% budżetu organizacji. Sytuacja taka nie jest satysfakcjonująca dla Polski. Jako że projekt LHC najpierw polega na wybudowaniu dużej infrastruktury badawczej, a później dopiero następuje eksploatacja, 15% budżetu projektu zostało dotychczas przeznaczone na usługi budowlane. Należy jednak podkreślić, że rynek CERN i LHC jest bardzo różnorodny. Dostawcami są przedsiębiorstwa z prawie wszystkich gałęzi przemysłu. Poczynając od dostaw produktów zaawansowanych technicznie, tj. nadprzewodzących magnesów, instalacji kriogenicznych, usług budowlanych, geologicznych, instalacyjnych, a kończąc na usługach związanych z zarządzaniem nieruchomościami czy sprzątaniem. Jeśli chodzi o udział Polski w dostawach, to przedsiębiorstwa polskie dostarczyły do CERN produkty i usługi o łącznej wartości przekraczającej 50 mln franków szwajcarskich. Polską specjalnością stały się instalacje gazowe ze stali szlachetnych i metali kolorowych (Przedsiębiorstwo ZEC Service), instalacje elektryczne (Energomontaż Południe), szafy sterownicze (ZPAS Przysiółko), elementy elektroniki przemysłowej (Zakłady Elektroniki Górniczej ZEG Tychy), elementy kriogeniczne. Niestety, bankructwo JTT wstrzymało dostawy serwerów z Polski. Dostawy do CERN realizowały między innymi RAFAKO z Raciborza (kriostaty), CHEMAR Kielce (stanowisko do testów prototypowych dipolowych magnesów nadprzewodzących) czy METALCHEM Kościan (dennice do zbiorników próżniowych). Dostawy były też realizowane przez polskie ośrodki naukowe, jak np. Instytut Fizyki Jądrowej

wej, Politechnika Wrocławska, Politechnika Krakowska. Należy podkreślić fakt, że dostawy z Polski są uznawane za spełniające wymagania technologiczne oraz organizacyjne. Polskie przedsiębiorstwa dzięki temu nabierają doświadczeń technologicznych i proceduralnych niezbędnych do stania się partnerami w innych dużych projektach badawczych [4]. Niemniej jednak ich udział w projekcie jest zbyt niski. Nawiasem mówiąc, nie tylko Polska ma z tym problem, cały projekt boryka się z problemem niezbilansowania części krajów członkowskich.

Jeśli chodzi o Polskę, to nie ma dużego zainteresowania ze strony polskich przedsiębiorstw udziałem w dostawach towarów i usług do CERN. Nie mają one świadomości, jak różnorodna jest skala możliwych dostaw i usług – poczynając od dostaw stali czy usług budowlanych, a kończąc na usługach sprzątnia czy usługach biur turystycznych. Niemalże każdy przedsiębiorca może znaleźć coś dla siebie. Polscy przedsiębiorcy mają obawy związane ze startem w przetargach międzynarodowych przede wszystkim ze względu na to, że oferta musi być przygotowana w języku angielskim (bariera językowa), a ponadto wymaga to kontaktu (często udziału w zagranicznych podróżach służbowych) z przedstawicielami CERN.

Aby przybliżyć polskim przedsiębiorcom i innym polskim jednostkom możliwości współpracy z CERN w ramach dostaw różnorodnych produktów i usług oraz aby aktywnie przyczynić się do promocji polskich dostawców, w 1995 roku zorganizowano wystawę „Poland at CERN 1995”, a w 2000 roku kolejną: „Poland at CERN 2000”. W tych przedsięwzięciach, zorganizowanych przez zespół prof. M. Chorowskiego pracujący pod auspicjami Państwowej Agencji Atomistyki, wzięły udział liczne przedsiębiorstwa z Polski, w tym także instytuty naukowo-badawcze. Dostawy do CERN w latach 1996 i 1997 utrzymały swój relatywnie wysoki poziom, po czym w kolejnych latach uległy tendencji spadkowej, by znów po 2000 roku zacząć rosnąć, co można uzasadnić oddziaływaniem wystaw. Na tej podstawie nasuwa się wniosek, iż przedsiębiorstwa w Polsce nie dysponują wiedzą dotyczącą szerokiego spektrum możliwości, jakie oferuje im projekt LHC. Brak odpowiedniej wiedzy powoduje, że przedsiębiorstwa, które spełniałyby wszelkie wymogi technologiczne oraz organizacyjne, nie decydują się na złożenie oferty. Do tego dochodzi bariera językowa oraz obowiązek częstych podróży służbowych związanych z utrzymaniem kontaktu z przedstawicielami CERN [3].

Niezbilansowanie polskiej składki członkowskiej wynika także z nieufności, jaka panowała w CERN wobec wykonawców z Polski podczas podejmowania decyzji o największych kontraktach LHC, a także z rozstrzygnięć przetargów, wyraźnie faworyzujących firmy lokalne (te zarejestrowane we Francji i Szwajcarii). Jedną z przyczyn może być tutaj duże rozproszenie geo-

graficzne przedsiębiorstw z państw mających „prawo do projektu”. Łatwiej pod względem logistycznym jest przeprowadzić audyt technologiczny i potwierdzić kompetencje przedsiębiorstw lokalnych niż tych rozproszonych w różnych krajach Europy. Ponadto, często mając wybór, mamy większe zaufanie do naszych lokalnych przedsiębiorców. Samym przedsiębiorstwom jest też z pewnością łatwiej utrzymywać kontakty z CERN, jeśli ich siedziba mieści się w tym samym kraju. Jednak możliwości komunikowania się czy podróżowania są bardzo duże w obecnych czasach, dlatego rozproszenie geograficzne dostawców nie powinno stanowić istotnej bariery. Poza tym w żaden sposób nie można usprawiedliwiać faworyzowania przedsiębiorstw lokalnych [3].

Udział w projekcie może być zyskowny, ale liczą się przecież nie tylko zyski ekonomiczne, ale sam prestiż związany z udziałem w takim projekcie. Przedsiębiorstwa powinny traktować udział w projekcie jako inwestycję i swoistą reklamę własnej działalności. Udział w takim projekcie jest bardzo ważnym doświadczeniem, które może pomóc w przyszłości wygrać inne przetargi. Rynek tworzony przez wielkie projekty badawcze jest rynkiem wysoce specjalistycznym, na którym istnieje duża bariera wejścia. Jest to rynek specyficzny, bo tworzony przez wiele krajów. Swoje produkty i usługi dostarczają najlepsze firmy z całego świata. Dla polskich przedsiębiorstw udział w takich projektach stanowiłby okazję do tego, by nawiązać kontakty oraz współpracę z wiodącymi, światowymi przedsiębiorstwami, które działają w obszarze high-tech [4]. Dlatego szkoda, iż szansa ta nie jest przez polskie przedsiębiorstwa wykorzystywana.

4. Ocena obecnego systemu wyboru dostawców w projekcie „Wielki Zderzacz Hadronów”

Budowa LHC wymagała opracowania nowego modelu współpracy pomiędzy dużym laboratorium badawczym a przemysłem. Wymagało to opanowania zdolności współpracy na niespotykaną dotąd skalę pomiędzy środowiskiem naukowym a środowiskiem przemysłowym. Doświadczenie to jest bardzo ważne, gdyż duże projekty badawcze, wymagające budowy dużych infrastruktur badawczych, nie skończą się w Europie na projekcie LHC. Tego typu projekty będą się pojawiały nieprzerwanie przez wiele lat. Dlatego też ważne jest opanowanie takiej umiejętności współpracy i gromadzenie wszelkich doświadczeń.

Procedury zakupów i wyłaniania dostawców przez CERN stanowią często wzorce lub punkty odniesienia dla innych tego typu projektów. Rozwiązania technologiczne zastosowane w LHC wymagały rozwoju i zastosowania nowych technologii, którymi przemysł wcześniej nie dysponował, oraz musiały industrializację i rozproszenie produkcji elementów i podzespołów

akceleratora, która była prowadzona we wszystkich krajach członkowskich CERN oraz w USA, Japonii, Rosji, Indiach. W projekcie LHC mamy do czynienia ze sprzężeniem zwrotnym pomiędzy CERN a przemysłem. Technologie były opracowywane w CERN, a następnie przemysł był angażowany w produkcję prototypów. Następnie rozwiązania były udostępniane podmiotom gospodarczym, które dokonywały wdrożenia często modyfikując koncepcje pierwotnie zaproponowane przez fizyków i inżynierów z CERN. Wielkość tego projektu spowodowała, że stał się on sam w sobie na tyle głębokim i stabilnym zaawansowanym technologicznie rynkiem, że wiele przedsiębiorstw dostrzegając korzyści ekonomiczne zdecydowało się na uruchomienie specjalnych linii technologicznych pracujących przez kilka lat tylko na potrzeby LHC. Poza tym okazało się, że te nowe rozwiązania i technologie mogą znaleźć swoje zastosowanie w innych dużych projektach oraz w innych dziedzinach życia np. w energetyce, informatyce czy medycynie [4].

Niemniej jednak każdy system decyzyjny może i powinien podlegać procesowi uczenia się, zwłaszcza jeśli łatwo w nim dostrzec pewne niedoskonałości. Naszym zdaniem system decyzyjny stosowany w projekcie LHC jest zbyt „twardy” – w tym sensie, że diametralnie zmienia (wyostrza) sposób postępowania po przekroczeniu ostrych granic cenowych, 10 000 CHF i 200 000 CHF. W przedziale spodziewanej ceny (10 000 CHF, 200 000 CHF) procedura postępowania jest zupełnie inna niż powyżej dolnej granicy tego przedziału, a powyżej obowiązuje jeszcze inna procedura. Jednak w samym przedziale (10 000 CHF, 200 000 CHF) jest ona taka sama, niezależnie od tego, w której jego części się znajdujemy. Nie jest to zgodne z intuicją: jeśli faktycznie cena decyduje o „ostrości” procedur wyboru, to kontrakt o spodziewanej cenie około 180 000 CHF powinien być traktowany inaczej niż kontrakt o spodziewanej cenie około 12 000 CHF. W literaturze problem wyboru dostawców jest rozwiązywany często za pomocą miękkich, elastycznych kryteriów [2], w omawianym projekcie nie zostało to zastosowane.

Kolejnym zarzutem jest to, iż wydaje się, że obecny system wyboru dostawców nie jest w rzeczywistości do końca taki, jak został zapisany na papierze. Zarzuty o stronniczość (o wybór dostawców lokalnych) czy uwagi o preferowaniu dostawców z krajów niezbilansowanych, a także raczej nieformalna ocena wynikająca z „wizji lokalnej”, choć mają różną wagę z etycznego punktu widzenia, świadczą razem właśnie o tym, iż poza zapisanymi regułami stosuje się jeszcze inne, i to miękkie, a na pewno niesformalizowane. Być może brana jest pod uwagę wiarygodność danego kraju, może wygoda – do tego zarządzający projektem mają prawo, ale powinno to zostać zapisane. Podobnie zapisany powinien być formalny wpływ kryterium braku zbilansowania. Znowu jest to kryterium w ewidentny sposób miękkie. Kraj może być bardzo niezbilansowany, średnio niezbilansowany i znikomo zbilansowany – w każdym z tych przypadków powinien być traktowany inaczej.

Trzecim zarzutem jest to, iż wydaje się, że nie wszystkie kryteria zostały w projekcie uwzględnione. Uderzający jest fakt, iż nie uwzględnia się kryteriów typu: „ważność danej dostawy dla projektu”, „potencjalny wpływ problemów z daną dostawą (jej terminem, jakością produktu) na termin realizacji projektu”, „ważność danej dostawy dla rozwoju nauki” itp. Przecież cena nie jest jedynym czy nawet często w ogóle nie jest miernikiem spełnienia tego typu kryteriów. Trzeba zauważyć, że są to wszystko kryteria miękkie.

Ponadto, w obecny system wyboru dostawców nie ma wbudowanego procesu uczenia się, a reguły decyzyjne, wyznaczane na początku przez eksperta, muszą być weryfikowane w praktyce, zwłaszcza w tego typu projektach, które trwają latami i procesem zmian reguł decyzyjnych trzeba zarządzać. Weryfikacja reguł decyzyjnych w procesie uczenia się jest wpisana w każdy system ekspercki [1].

Zatem proponujemy zastosowanie w omawianym projekcie, jak i w innych projektach tego typu, rozmytego systemu wyboru dostawców, systemu dynamicznego, którego rozmyte reguły decyzyjne, jak i słownik, którym one się posługują, będą podlegały weryfikacji i modyfikacjom na podstawie konkretnych doświadczeń.

5. Idea rozmytego systemu wyboru dostawców w projekcie „Wielki Zderzacz Hadronów”

Zarówno w obecnym systemie wyboru dostawców, jak i w proponowanym, stosowane są i powinny być reguły decyzyjne, zdefiniowane dla N atrybutów decyzji A_i ($i = 1, \dots, N$) oraz M procedur postępowania P_i ($i = 1, \dots, M$). Każdy atrybut decyzji może przyjmować wartość liczbową lub lingwistyczną. Zbiór możliwych wartości przyjmowanych przez atrybut A_i oznaczmy jako Z_i ($i = 1, \dots, N$). Z_i ($i = 1, \dots, N$) mogą być albo podzbiorami liczb rzeczywistych, albo skończonymi zbiorami wartości lingwistycznych, jakie może przyjmować dany atrybut. W tym ostatnim przypadku Z_i jest językiem, skończonym zbiorem pojęć, które muszą przedtem zostać zdefiniowane.

Każda z procedur postępowania opisuje to, co należy zrobić w procesie wyboru dostawcy. W obecnym systemie mamy w sposób jawny wymieniony jeden atrybut – cenę. Atrybut ten przyjmuje wartości liczbowe. W sposób niejawni występują takie atrybuty, jak zbilansowanie kraju, wynik „wizji lokalnej”, ukryte preferencje decydentów itp. W sformalizowanym systemie te atrybuty musiałyby mieć swój jasno zdefiniowany język. Procedury postępowania w obecnym systemie to np. „Wybierz dostawcę proponującego najmniejszą cenę”, „Sprawdź wiarygodność kandydata na miejscu” itp.

Proponujemy zachować stosowaną formę procedur, ewentualnie rozszerzając ich zbiór o inne, ale jeśli chodzi o atrybuty, proponujemy, by zbiory Z_i ($i=1, \dots, N$) były wyłącznie skończonymi zbiorami wyrażeń lingwistycznych. Mielibyśmy zatem $Z_i = \{w_{i,1}, w_{i,2}, \dots, w_{i,n_i}\}$ ($i=1, \dots, N$), gdzie $w_{i,1}, w_{i,2}, \dots, w_{i,n_i}$ byłyby wyrażeniami lingwistycznymi. W ramach danego języka, czyli zbioru $Z_i = \{w_{i,1}, w_{i,2}, \dots, w_{i,n_i}\}$, muszą być ujęte negacje, koniunkcje i alternatywy pojęć podstawowych, np. jeśli mamy zbiór $Z_i = \{\text{małaśredni}, \text{duży}\}$, to musi on zostać uzupełniony o elementy „nieduży”, „niemały”, „mały lub duży” itp. Szczegóły dotyczące tego, w jaki sposób należy dokonywać uzupełnienia zbioru Z_i można znaleźć w literaturze [6].

Reguły decyzyjne systemu R_l ($l=1, \dots, T$) powinny być określone następująco:

1. Dla każdego $l=1, \dots, T$ jest zdefiniowany zbiór indeksów atrybutów występujących w danej regule, czyli $S_l \subset \{1, \dots, N\}$.

2. Dla każdego $t \in S_l$ jest zdefiniowana wartość tego atrybutu brana pod uwagę w danej regule, czyli w_{t,h_l} .

3. Dla każdego $l=1, \dots, T$ jest zdefiniowany zbiór indeksów procedur występujących w danej regule, czyli $Q_l \subset \{1, \dots, M\}$.

4. Sama reguła R_l ($l=1, \dots, T$) miałaby postać:

jeśli A_t jest w_{t,h_l} dla każdego $t \in S_l$, stosuj procedury P_m , $m \in Q_l$.

Zakładamy przy tym, że dla każdej pary różnych liczb $l_1, l_2 \in \{1, \dots, T\}$ mamy $S_{l_1} \neq S_{l_2}$ lub istnieje $t_0 \in S_{l_1} = S_{l_2}$ takie, że $w_{t_0,h_{l_1}} \neq w_{t_0,h_{l_2}}$. Warunek ten oznacza, że żadne dwie reguły nie mają takiej samej przesłanki. Jeśli w procesie generowania reguł tak się zdarzy, trzeba oszacować ważność każdej z reguł i wybrać tę ważniejszą [9].

Zbiory atrybutów, reguł, procedur, a także możliwych wartości atrybutów oraz ich znaczenia powinny być określone przez eksperta, a następnie powinny podlegać weryfikacji w trakcie realizacji projektu oraz dopasowaniu do specyfikacji konkretnych projektów. Atrybuty A_i ($i=1, \dots, N$) to np. wielkość ceny, ważność decyzji dla projektu, ważność decyzji dla rozwoju nauki, potencjalne konsekwencje nieterminowej dostawy, potencjalne konsekwencje złej jakości produktu, stopień niezbilansowania kraju, stopień wiarygodności kraju, ogólne wrażenie po wizji lokalnej, waga bliskości geograficznej itd. Odpowiadające im słowniki (Z_i ($i=1, \dots, N$)) to np. zbiory {znikoma, mała, średnia, duża bardzo duża}, {bardzo złe, złe, średnie, dość dobre, bardzo dobre, doskonałe}, {nieznaczące, istotne, poważne, ogromne} itd. Procedury P_j ($j=1, \dots, M$) to np.

„wybierz dostawcę proponującego najniższą cenę”, „nadaj rankingi dostawcom wg takiego a takiego wzoru i wybierz dostawcę o najwyższej ocenie”, „przeprowadź wizję lokalną”

Wartości lingwistyczne muszą być za każdym razem konkretnie zdefiniowane. Będą one zbiorami rozmytymi [5], co sprowadza się do tego, że definiowane pojęcia będą mogły mieć nieostre granice, a konkretne cechy mogą występować w różnych stopniach, od 0 (dana cecha nie występuje wcale) do 1 (dana cecha występuje w pełni), np. „duża cena to cena powyżej 200 000, cena poniżej 30 000 nie jest w żadnym stopniu duża, ale cena c z przedziału [30 000, 200 000] jest duża w stopniu $(c - 30000)/(200000 - 30000)$. Słowniki będą musiały być dopasowywane do sytuacji – w jednym projekcie znaczenie słowa „duża cena” może być zupełnie inne niż w drugim projekcie. Niektóre pojęcia ze słowników nie będą mogły być zdefiniowane za pomocą wzorów (np. co to znaczy, że podczas wizji lokalnej firma wywarła średnie wrażenie?), a będą raczej subiektywne, ale też powinny być tak opisane, by każdy ekspert posługiwał się tym samym językiem, dopasowanym do aktualnej sytuacji.

Reguły, tak jak wspomniano, mogą być modyfikowane. Mogą być też dodawane nowe reguły. Ważne, by system wyboru dostawcy był jak najbardziej skuteczny i uwzględniał wszystkie kryteria w sposób zgodny z ich znaczeniem dla celów projektu.

Za przykład reguł niech służą dwie następujące:

1. Jeśli znaczenie dostawy dla projektu duże i wiarygodność kraju mała, nie wybieraj danego dostawcy.
2. Jeśli potencjalne konsekwencje opóźnienia małe, potencjalne konsekwencje złej jakości małe, cena nie bardzo duża i kraj zbilansowany w stopniu małym, wybierz danego dostawcę.

Podsumowanie

Na gruncie analizy jednego z wielkich projektów badawczych zaproponowano koncepcję wyboru dostawców. Jest to bardzo istotny problem, z natury swojej wielokryterialny. Każdy problem wielokryterialny można rozwiązywać w różny sposób. Chodziło tu o to, by zaproponowany sposób był kontynuacją i uogólnieniem istniejących procedur, a także był prosty w zastosowaniu i opierał się na języku niezmatematyzowanym. Ponadto, musi być elastyczny i przejrzysty oraz musi podlegać procesowi ciągłej modyfikacji. Jest to bardzo istotne, gdyż gra toczy się często o duże pieniądze i bardzo istotne jest, by po pierwsze, podejmować decyzje właściwe w danym kontekście, a po drugie, by decyzje te były jasne, zgodne z celami projektu i nie wzbudzały protestów czy podejrzeń o stronniczość.

Oczywiście należy uwzględnić pytanie, czy proponowana koncepcja systemu może i powinna być stosowana w dużych projektach badawczych oraz jeśli tak, czy są szanse na jej akceptację ze strony wszystkich udziałowców projektu. W projektach finansowanych ze środków publicznych trzeba być szczególnie uczulonym na uniknięcie wszelkich zarzutów o niejasne kryteria czy manipulację w podejmowaniu decyzji, a na pierwszy rzut oka proponowany system może sprawiać wrażenie podatnego na tego typu podejrzania. Niemniej jednak w pierwszym rzędzie trzeba stwierdzić, że duże projekty badawcze charakteryzują się często kilkudziesięcioletnim okresem trwania i przy tak długim horyzoncie czasowym kryteria podejmowania decyzji w sprawie zakupów i tak podlegają zmianom. Tak było w przypadku omawianego tu projektu. Wskazane jest, by te zmiany – i tak nieuniknione, gdyż w ciągu kilkudziesięciu lat kontekst projektu z pewnością ulegnie zmianie – podlegały zarządzaniu (zarządzanie zmianami w projekcie jest nieodzowną częścią zarządzania projektem), a dobre zarządzanie musi – zwłaszcza w przypadku tak długich projektów – obejmować proces uczenia się. Celem przedstawionej tu propozycji jest zaproponowanie systemu zarządzania zmianami w systemie zakupów w dużych projektach badawczych, właśnie po to, by zmiany te podlegały planowaniu i kontroli wszystkich udziałowców projektu.

Drugim aspektem, jaki warto podkreślić jest to, że „system rozmyty” nie znaczy „system nieprecyzyjny”. Rozmyte systemy decyzyjne pozwalają uwzględnić różne kryteria w różnym stopniu, ale sposób tego uwzględnienia, czyli konkretny wzór rangujący potencjalnych dostawców, jest z góry określony i jednoznaczny. Musi być oczywiście upubliczniony odpowiednio wcześniej, przed ogłoszeniem kolejnych przetargów i dokładnie wyjaśniony. Ponieważ będzie to zawsze jeden wzór, wynikający z przyjętej wagi różnych kryteriów, jego zastosowanie może być u potencjalnych dostawców zautomatyzowane, choćby za pomocą tak popularnego narzędzia, jakim jest EXCEL. Stąd wynika prostota w jego zastosowaniu – potencjalny dostawca będzie mógł wprowadzać różne warianty swojej oferty i uzyskać od razu jednoznaczną ocenę liczbową swojej propozycji, zgodną z aktualnymi preferencjami kierownictwa czy sponsorów projektu. Wydaje się zatem, iż są szanse na zastosowanie proponowanego systemu w praktyce. Szanse te muszą być oczywiście zweryfikowane w praktyce, co jest kolejnym zadaniem badawczym zaplanowanym do realizacji. Dalszym krokiem będzie próba weryfikacji zaproponowanej koncepcji w projekcie „Wielki Zderzacz Hadronów”, w którym, jak wspomniano, uczestniczy Politechnika Wrocławska.

Literatura

1. Barrenechea, Edurne, Bustince H., Calderon M., Fernandez A., Herrera F. (2010). Solving Multi-class Problems with Linguistic Fuzzy Rule Based Classification Systems Based on Pairwise Learning and Preference Relations. *Fuzzy Sets and Systems*, 161, 23, Theme: Information Processing, s. 3064-3080.
2. Batieha F., Dalalah, D., Hayajneh M. (2011). A Fuzzy Multi-criteria Decision Making Model for Supplier Selection. *Expert Systems with Applications*, 38, 7, s. 8384-8391.
3. Chorowski M. (2010). Syntetyczna ocena udziału polskiego przemysłu w budowie infrastruktury badawczej w CERN w latach 1991-2009. Wrocław.
4. Foresight dla energetyki termojądrowej (opracowania i ekspertyzy). (2009). Red. Z. Składanowski, A. Marchewka. Instytut Fizyki Plazmy i Laserowej Mikro-syntezy, Warszawa.
5. Kuchta D. (2001). Miękka matematyka w zarządzaniu. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
6. Computing with Words in Information/intelligent Systems. (1999). Red. L.A. Zadeh, J. Kacprzyk. Physica-Verlag, Heidelberg.
7. Witryna internetowa Technology Transfer Agency Techtra Sp. z o.o. www.techtra.pl
8. Witryna Internetowa Europejskiej Organizacji Badań Jądrowych (CERN) www.cern.ch
9. Yen V.C. (1999). Rule Selections in Fuzzy Expert Systems. *Expert Systems with Applications*, 16, s. 79-84.

PURCHASE PROBLEM IN BIG RESEARCH PROJECTS

Summary

In large scale research projects (projects that the time of their realization takes years and their budgets are counted in billions euro) there are usually a lot of different type of deliveries (some elements are cheap, some are really expensive, some of them are maybe cheap but very important for project, etc.). Responsible people in project team are made a lot of important decisions connected with choosing right supplier among huge number of different suppliers. Potential suppliers are very often unknown and their seats are situated in many places all over the world that's why many decisions which are made are very risky. The biggest risk is connected with delay in supplies and/or with quality of deliveries. In this article, based on real project called Large Hadron Collider will be proposed conception of choosing suppliers or making decision about better check (or verification) of quality of potential suppliers (this is usually connected with travel and long stay in potential suppliers seats to make technological audit). In this conception will be used fuzzy number.

Anna Ślusarczyk

Politechnika Wrocławska

Dorota Kuchta

Politechnika Wrocławska

Wyższa Szkoła Oficerska Wojsk Lądowych im. gen. Tadeusza Kościuszki we Wrocławiu

ZARZĄDZANIE RYZYKIEM OPÓŹNIENIA PROJEKTÓW REALIZOWANYCH PRZEZ JEDNOSTKI SAMORZĄDOWE O PROSTEJ STRUKTURZE SIECI PROJEKTOWEJ

Wprowadzenie

Jednostki samorządowe w Polsce realizują liczne projekty, które wzbudzają nieraz spore emocje w społeczeństwie. Ludzie mają bowiem poczucie, iż przynajmniej częściowo i oni za nie płacą, a często są i bezpośrednio dotknięci przez sam proces realizacji tych projektów. Dotyczy to zarówno „spektakularnych” projektów inwestycyjno-budowlanych, jak i projektów wdrożeniowych, kiedy w urzędach wprowadzany jest nowy system obsługi czy inne zmiany, które mają pomóc społeczeństwu, ale społeczeństwo niekoniecznie postrzega ten proces pozytywnie. Dzieje się tak zwłaszcza w przypadku opóźnień, które niestety są – ogólnie w projektach niemal każdej branży – czymś niemal normalnym. Jeśli nie ma opóźnień, to często okazuje się, że oddano produkt niedokończony i tak naprawdę opóźnienie – w sensie terminu oddania produktu dobrego i całego – jednak jest. Ponadto, w przypadku projektów współfinansowanych przez Unię Europejską trzeba trzymać się sztywnych terminów. Dlatego w tego typu projektach bardzo ważne jest efektywne zarządzanie ryzykiem opóźnienia projektów. Musi ono być skuteczne, a jednocześnie proste.

Efektywne zarządzanie ryzykiem opóźnienia projektu, zwłaszcza w przypadku projektów, w których występuje wielu wykonawców i wielu udziałowców (a tak jest w projektach realizowanych przez urzędy), musi uwz-

ględniać strukturę sieci projektowej. To ona bowiem pozwala określić wpływ poszczególnych zadań czy części projektu na ryzyko opóźnienia całego projektu, co jest od dawna oczywiste (powszechnie stosowana metoda ścieżki krytycznej oparta jest właśnie na strukturze sieci). Inne metody zarządzania ryzykiem, takie jak HAZOP FMEA, listy kontrolne, wspomagają zarządzanie ryzykiem opóźnienia projektu, ale w ostatecznym rozrachunku powinny być odniesione do sieci projektu i wpływu poszczególnych jej części na opóźnienie projektu.

Proponowane w literaturze metody analizy sieci projektu pod kątem ryzyka opóźnienia projektu są zazwyczaj „naukowe” (oparte na teorii prawdopodobieństwa), zatem zbyt skomplikowane z punktu widzenia praktyków realizujących projekty w urzędach państwowych, ale jednocześnie w pewnym sensie uproszczone, bo oparte na licznych założeniach. Z jednej strony modeluje się niepewne czasy trwania czy koszty poszczególnych zadań, reprezentując je za pomocą zmiennych losowych przedziałów [2] czy liczb rozmytych [4], przyznając, że na etapie planowania projektu nie potrafimy ich zazwyczaj oszacować dokładnie, a w następujących potem obliczeniach zakłada się zazwyczaj niezależność poszczególnych zmiennych od siebie, co jest założeniem w praktyce nieprawdziwym. Ponadto, mając na celu oszacowanie czasów trwania projektu, wyznacza się ścieżki krytyczne (oparte na różnie definiowanych wartościach średnich – oczekiwany – liczb rozmytych czy zmiennych losowych) bądź ścieżki o wysokim stopniu krytyczności [4] i na tym analiza się kończy. Oczywiście informacja o tym, która ścieżka będzie w średnim przypadku krytyczna (tak robi się w powszechnie znanej metodzie PERT) bądź które ścieżki mogą być krytyczne z dużym stopniem możliwości jest ważna, ale może nie być do końca adekwatna, jeśli założenie o niezależności czasów trwania czy kosztów poszczególnych zadań nie jest spełnione (a w praktyce bardzo często nie jest). Ponadto, przypadek średni jest tylko przypadkiem średnim i tak naprawdę niekoniecznie ma on praktyczne znaczenie, bo niekoniecznie wystąpi. Trzeba też pamiętać o tym, że oszacowanie rozmytych czy losowych liczb reprezentujących parametry poszczególnych zadań może być niezwykle trudne, zwłaszcza jeśli dużo różnych czynników decyduje o niepewności w ich oszacowaniu (czynniki ludzkie wewnętrzne, otoczenie projektu, czynniki techniczne, pogodowe itd.). W rezultacie stosuje się probabilistyczne czy rozmyte oszacowanie reprezentujące „mieszankę” wpływu różnych czynników losowych. Jeśli w praktyce problemy pojawią się w jednym z wymienionych aspektów, to trudno będzie „wyluskać” wpływ tego właśnie aspektu na ryzyko opóźnienia projektu.

W przypadku większości projektów (jeśli pojęcie „większości” odniesiemy do teoretycznego zbioru wszystkich możliwych sieci projektowych o ustalonej maksymalnej liczbie zadań) inne podejście nie jest możliwe z powodu złożoności problemu, chyba że zrezygnujemy z faktycznie efektywnego i skutecznego zarządzania ryzykiem opóźnienia projektu. Nawet sieć o stosunkowo niewielkiej liczbie zadań może mieć dużo różnych ścieżek, a także dużo różnych zależności między zadaniami, których uwzględnienie *explicite* byłoby praktycznie niemożliwe. Trzeba więc przyjmować upraszczające założenia, żeby uzyskać jakieś orientacyjne informacje na temat projektu i jego najbardziej newralgicznych punktów. Niemniej jednak należy postawić pytanie, czy zbiór sieci projektów występujących w praktyce, przynajmniej w niektórych dziedzinach, nie jest zbiorem znacznie różniącym się od teoretycznie generowanego zbioru wszystkich możliwych sieci projektowych o ustalonej maksymalnej liczbie zadań. Doświadczenie w zakresie projektów realizowanych przez urzędy pozwala postawić hipotezę, iż te zbiory bardzo się od siebie różnią pod względem złożoności struktury sieci. Wydaje się, iż w praktyce występują często sieci o tak prostej strukturze, iż ich analiza może być znacznie dokładniejsza i lepiej odwzorowywać rzeczywistość.

Literatura pokazuje wyraźnie, iż tzw. morfologia sieci ma znaczny wpływ na ryzyko opóźnienia projektu, a także na możliwości zarządzania tym ryzykiem. Tavares inni [14] proponują wskaźniki określające morfologię sieci i wykazują korelację między ryzykiem opóźnienia projektu a wartością odpowiednich wskaźników. Natomiast Zammori inni [14] pokazują, że w sieciach o nieskomplikowanej morfologii można ryzyko opóźnienia kontrolować znacznie efektywniej niż w ogólnym przypadku, rezygnując z upraszczających i nie-raz zniekształcających rzeczywistość założeń. W rozpatrywanym studium przypadku, jak się wydaje reprezentatywnym dla pewnej grupy projektów realizowanych przez jednostki samorządowe, struktura sieci jest na tyle nieskomplikowana, iż wydaje się, że jest możliwa dość prosta, a jednocześnie skuteczna analiza ryzyka opóźnienia projektu. Głównym celem opracowania jest zaprezentowanie jej idei.

Struktura artykułu jest następująca: najpierw przedstawimy zarys znanej z literatury analizy wpływu struktury sieci na ryzyko opóźnienia projektu i na stopień trudności zarządzania ryzykiem opóźnienia projektu. Następnie zostanie opisany badany projekt, realizowany przez wrocławskie starostwo powiatowe. Wreszcie zaproponujemy metodę zarządzania ryzykiem opóźnienia projektu dedykowaną przedsięwzięciom o prostej strukturze sieci projektowej.

1. Analiza ryzyka opóźnienia projektów przy różnych typach struktury sieci

W ogólnym przypadku istnieją różne metody analizy ryzyka opóźnienia projektów [11; 12]. Niemniej jednak są one zazwyczaj na tyle „zmatematyzowane”, iż nie będą przydatne w codziennej praktyce jednostek samorządowych. Część z tych metod opiera się na podejściu probabilistycznym – w tym przypadku obliczane są wartości średnie lub rozkłady planowanego czasu zakończenia projektu, natomiast ścieżka krytyczna wyznaczana jest na podstawie wartości oczekiwanych czasów trwania poszczególnych zadań projektu. Inne metody opierają się na podejściu przedziałowym lub rozmytym [1; 3; 4]. Również w tych przypadkach wyznacza się średni albo rozmyty czas trwania projektu, a ścieżki krytyczne wyznaczane są albo na podstawie średnich czasów trwania* albo dla różnych ścieżek wyznacza się stopnie ich krytyczności i rozpatruje się ścieżki najbardziej krytyczne. Inne sposoby zarządzania ryzykiem opóźnienia projektów opierają się na stosowaniu buforów w różnych miejscach sieci [5; 9], niemniej jednak sposób ich rozmieszczenia, a także wyznaczanie ich wielkości są jeszcze zbyt mało rozwinięte z praktycznego punktu widzenia, by mogły być stosowane w urzędach.

W kilku artykułach zwraca się uwagę na wpływ struktury sieci na analizę ryzyka opóźnienia. Struktura sieci okazuje się mieć wpływ na złożoność obliczeniową problemu wyznaczania stopnia krytyczności poszczególnych ścieżek w przypadku rozmytych czasów trwania zadań [3]. Ponadto, struktura sieci okazuje się mieć wpływ na prawdopodobieństwo opóźnienia projektu. Pokazano [13], iż wskaźniki, takie jak liczba ścieżek w stosunku do liczby zadań czy maksymalna liczba relacji, jakimi powiązane jest jedno zadanie z innymi, czy też liczba ścieżek częściowych i ich powiązań z innym ścieżkami ma znaczny wpływ na prawdopodobieństwo opóźnienia projektu.

Zammori i inni [14] zajmują się sieciami projektowymi, przy których liczba wszystkich ścieżek od początku do końca projektu jest na tyle mała, iż każdą ze ścieżek można zająć się osobno. Autorzy proponują rozpatrywać każdą ścieżkę projektu osobno i ocenić jej „szanse” na stanie się ścieżką krytyczną podczas realizacji projektu z różnych punktów widzenia. Stosują między innymi następujące kryteria oceny stopnia krytyczności ścieżki:

* Wartość średnia liczby rozmytej może być definiowane na różne sposoby [8].

a. Możliwa zmienność długości zadań na ścieżce wynikająca z wewnętrznych uwarunkowań organizacji, z „normalnej” zmienności warunków realizacji zadania (gorsze lub lepsze samopoczucie wykonawców, większe lub mniejsze natężenie normalnych obowiązków pozaprojektowych, codzienne problemy i zaburzenia w pracy, niewiedza czy brak umiejętności/doświadczenia członków organizacji).

b. Potencjalne możliwości skrócenia pewnych zadań na ścieżce w razie potrzeby (czyli w razie przedłużenia się innych zadań na ścieżce bądź innych ścieżek), oczywiście związane z dodatkowymi kosztami – ten czynnik, w przeciwieństwie do pozostałych, powinien być jak największy.

e. Ryzyko znacznych poprawek – chodzi o identyfikację tych zadań na ścieżce bądź wpływających na zadania na ścieżce, po wykonaniu których może się okazać, iż zadanie trzeba wykonać jeszcze raz. Są to zazwyczaj zadania o dużym stopniu nowości dla danej organizacji. W przypadku problemów podczas realizacji takich zadań osoby odpowiedzialne za daną ścieżkę i za ścieżki uzależnione od tego zadania powinna być powiadomione, żeby odpowiednio zmienić oszacowanie czasu realizacji całych ścieżek.

d. Ryzyko opóźnień wynikające z dzielenia zasobów z innymi zadaniami danego projektu, ale przede wszystkim z innymi projektami. Osoba odpowiedzialna za daną ścieżkę powinna takie zadania zidentyfikować i być uwrażliwiona na problemy w zadaniach/projektach, z którymi zadania „jej” ścieżki dzielą zasoby.

e. Możliwa zmienność długości zadań na ścieżce wynikająca z wpływu czynników zewnętrznych. Osoba odpowiedzialna za daną ścieżkę powinna zidentyfikować te zadania, na których czas trwania mogą mieć wpływ zdarzenia niezależne od organizacji i być uwrażliwiona na niebezpieczeństwo materializacji tego typu zdarzeń.

Zammori i inni [14] stosują aparat liczb rozmytych do kwantyfikacji poszczególnych kryteriów, a następnie jedną z metod rozmytej oceny wielokryterialnej, by ocenić stopień krytyczności poszczególnych ścieżek. Jednak ze względu na adresata systemu zarządzania ryzykiem, jakim w przypadku niniejszej propozycji są jednostki samorządowe, w których tego rodzaju podejścia nie mają szans na praktyczne zastosowanie [10], rezygnujemy z tej części propozycji autorów artykułu, ograniczając się do określenia stopnia spełnienia kryteriów a-e. Resztę pozostawiamy nieformalnej ocenie osób odpowiedzialnych za poszczególne ścieżki. Podejście to zilustrujemy na przykładzie studium.

2. Opis projektu

Niniejszy i kolejny podrozdział został opracowany na podstawie dokumentów wewnętrznych i wywiadów z pracownikami Starostwa Powiatowego we Wrocławiu. Trzeba zaznaczyć, że dokumentacja była bardzo niekompletna, stąd pojawiła się konieczność opracowywania post factum niektórych dokumentów, które zgodnie z zasadami zarządzania projektami powinny były powstać na etapie planowania projektu.

Przedstawiany projekt znany jest pod nazwą „Rozbudowa infrastruktury teleinformatycznej na obszarze powiatu wrocławskiego i 7 gmin oraz wprowadzenie i zwiększenie dostępności elektronicznych usług dla mieszkańców i podmiotów gospodarczych regionu powiatu i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka, Żórawina” [6].

Głównym celem projektu jest uzupełnienie i poprawa funkcjonalności infrastruktury teleinformatycznej oraz budowa zintegrowanego systemu informatycznego wspomagających zarządzanie współpracujących jednostek tj. Powiatu Wrocławskiego i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka, Żórawina, a także zwiększenie dostępności do administracyjnych usług elektronicznych dla mieszkańców Powiatu Wrocławskiego.

Cele szczegółowe zostały określone w sposób następujący:

Wzrost jakości i szybkości obsługi mieszkańców i podmiotów gospodarczych.

Racjonalizacja wydatków urzędów, zwiększenie efektywności zarządzania jednostką.

Stworzenie usług elektronicznych dla mieszkańców powiatu i inwestorów.

Usprawnienie obiegu dokumentacji dzięki wprowadzeniu elektronicznego obiegu dokumentów.

Ułatwienie archiwizacji danych – forma elektroniczna.

Ułatwienie kontaktu z urzędami oraz uzyskania informacji jawnych.

Ułatwienie współpracy i komunikacji pomiędzy urzędami i jednostkami pomocniczymi.

Projekt zakłada między innymi zakup komputerów i oprogramowania, mający na celu podniesienie efektywności pracy urzędników powiatu wrocławskiego i gmin partnerskich, zwiększając tym samym standard usług ofero-

wanych klientom. Wdrożone rozwiązania umożliwią mieszkańcom na terenie powiatu wrocławskiego załatwianie określonych spraw urzędowych on-line, poprzez dostęp do elektronicznych skrzynek podawczych w urzędach.

Ponadto, utworzenie e-punktów w sołectwach oraz dwóch Biur Obsługi Klienta w Smolcu i Gniechowicach umożliwi korzystanie z tego serwisu tym klientom, którzy nie mają stałego dostępu do Internetu.

Zakres prac projektowych został podzielony na fazy główne – Przygotowanie dokumentacji i organizacja projektu, Przygotowanie infrastruktury – sprzęt, Zakup Platformy Transakcyjnej(PT), E-obługa interesantów, Zewnętrzny audyt projektu, Pozostałe wdrożenia. Następnie, zgodnie z logiką tworzenia WBS, nastąpiła ich dekompozycja na zadania bardziej szczegółowe (tabela 1).

Tabela 1

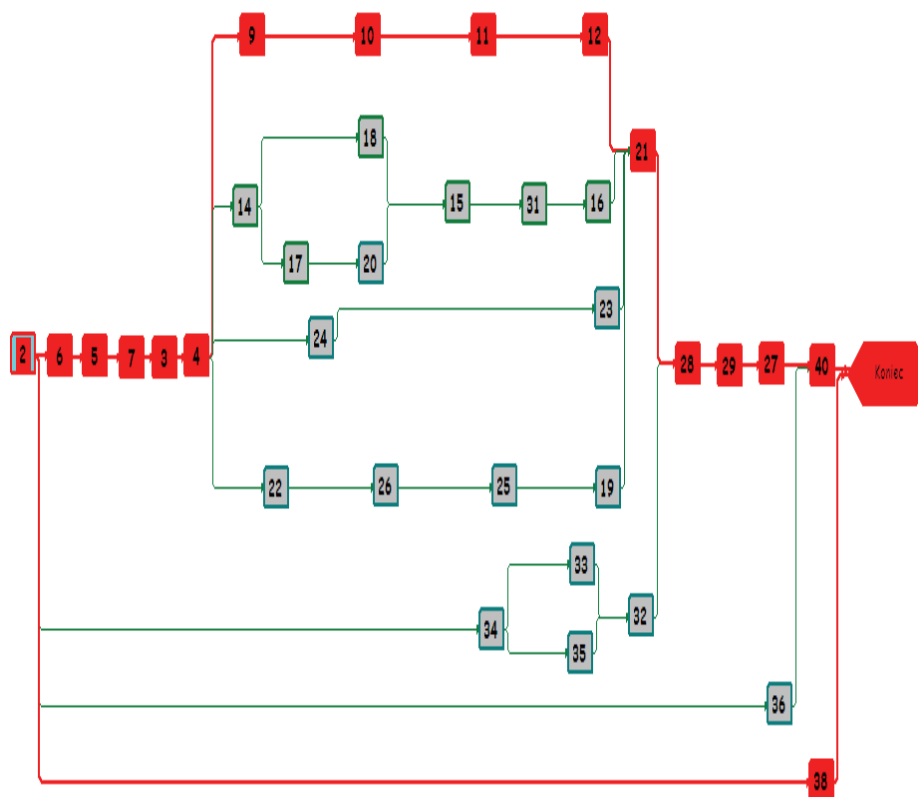
Struktura podziału pracy dla realizacji projektu „Rozbudowa infrastruktury teleinformatycznej na obszarze powiatu wrocławskiego i 7 gmin oraz wprowadzenie i zwiększenie dostępności elektronicznych usług dla mieszkańców i podmiotów gospodarczych regionu powiatu i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka, Żórawina”

Nr zadania	Nazwa zadania
1	2
1	Przygotowanie dokumentacji i organizacja projektu
2	Podpisanie porozumienia z partnerami
3	Opracowanie studium wykonalności
4	Złożenie dokumentacji projektowej wraz z wnioskiem o dofinansowanie z UE
5	Stworzenie zespołu ds. projektu
6	Zewnętrzne zarządzanie projektem
7	Szczegółowa specyfikacja wymagań
8	Przygotowanie infrastruktury – sprzęt
9	Rozbudowa połączeń sieciowych
10	Zakup sprzętu teleinformatycznego
11	Przydział zakupionego sprzętu teleinformatycznego
12	Instalacja sprzętu teleinformatycznego
13	Dostawa, instalacja i wdrożenie oprogramowania platformy teleinformatycznej e-Urząd
14	Zakup Platformy Transakcyjnej (PT)
15	Integracja poszczególnych obszarów funkcjonalnych platformy transakcyjnej
16	Integracja systemów wewnętrznych zarządzania zadaniami jednostek z PT
17	22 bazy danych czterech typów

cd. tabeli 1

1	2
18	Elektroniczne formularze
19	Utworzenie Lokalnego Centrum Certyfikacji
20	Biuletyn Informacji Publicznej – rejestry on-line
21	Integracja platformy transakcyjnej, BIP, MZZL oraz serwisu www SPW
22	Wybór administratorów systemu
23	System rejestracji czasu pracy
24	Prace nad Modułem Zarządzania Zasobami Ludzkimi (MZZL)
25	Zakup elektronicznego podpisu kwalifikowanego
26	"System Bezpieczeństwa związany z wdrożeniem ""e-Urząd""
27	Pilotażowe wdrożenie platformy transakcyjnej
28	Konfiguracja i uruchomienie sprzętu i oprogramowania we wszystkich lokalizacjach projektu
29	Szkolenia z zakresu obsługi aplikacji
30	E-obsluga interesantów
31	Stworzenie jednolitego interfejsu użytkownika
32	Aktualizacja informacji na serwerze www
33	Organizacja usługi telefonicznego wsparcia dla użytkowników
34	Utworzenie Biura Obsługi Klienta
35	Utworzenie nowych e-punktów kontaktu klienta z urzędem
36	Przeprowadzenie akcji informacyjno-promocyjnej projektu
37	Zewnętrzny audyt projektu
38	Zewnętrzny audyt projektu
39	Pozostałe wdrożenia
40	Pozostałe wdrożenia
41	Koniec

Jednym z etapów prac nad harmonogramem niniejszego projektu były rozmowy z pracownikami starostwa powiatu wrocławskiego, z osobami odpowiedzialnymi za zewnętrzne zarządzanie projektem oraz z podwykonawcami, w celu zidentyfikowania zależności pomiędzy zadaniami drugiego poziomu WBS. Następnie niezależny ekspert (niezaangażowany w pracę nad niniejszym projektem), posiadający ogromne doświadczenie w zakresie zarządzania projektami o tożsamym zakresie, dokonał weryfikacji wskazanych zależności. Te dwa źródła informacji zostały wykorzystane do stworzenia (post factum) harmonogramu realizacji projektu, który w ujęciu sieci projektu prezentowany jest poniżej (rys. 1).



Rys. 1. Sieć projektu „Rozbudowa infrastruktury teleinformatycznej na obszarze powiatu wrocławskiego i 7 gmin oraz wprowadzenie i zwiększenie dostępności elektronicznych usług dla mieszkańców i podmiotów gospodarczych regionu powiatu i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka, Żórawina”

Sieć projektu okazała się siecią o relatywnie prostej strukturze. Ścieżki mają części wspólne, ale tylko na swoich początkach i końcach. Zadania zaznaczone na rys. 1 ciemniejszym kolorem są krytyczne w przypadku oszacowań czasów trwania zadań przyjętych na etapie planowania projektu (do ich wyznaczania zastosowano klasyczną metodę ścieżki krytycznej).

3. Zarządzanie ryzykiem opóźnienia projektu w powiązaniu ze specyficzną strukturą sieci projektowej

Wskazane zależności pomiędzy zadaniami zdecydowały o kształcie sieci projektu, w której można wyróżnić dziewięć ścieżek (tabela 2).

Tabela 2

Identyfikacja ścieżek w sieci projektu „Rozbudowa infrastruktury teleinformatycznej na obszarze powiatu wrocławskiego i 7 gmin oraz wprowadzenie i zwiększenie dostępności elektronicznych usług dla mieszkańców i podmiotów gospodarczych regionu powiatu i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka, Żórawina”

Lp.	Ścieżka
I	2-6-5-7-3-4-9-10-11-12-21-28-29-27-40-41
II	2-6-5-7-3-4-14-18-15-31-16-21-28-29-27-40-41
III	2-6-5-7-3-4-14-17-20-15-31-16-21-28-29-27-40-41
IV	2-6-5-7-3-4-24-23-21-28-29-27-40-41
V	2-6-5-7-3-4-22-26-25-19-21-28-29-27-40-41
VI	2-6-5-7-3-4-34-33-32-28-29-27-40-41
VII	2-6-5-7-3-4-34-35-32-28-29-27-40-41
VIII	2-36-40-41
IX	2-38-41

Ścieżki o numerze jeden oraz dziewięć są ścieżkami krytycznymi. Niemniej jednak ścieżki te zostały uznane za krytyczne przy użyciu metody klasycznej i statycznej, nieuwzględniającej potencjalnej zmienności czasów trwania pozostałych zadań. Ryzyko zmiany czasu trwania pozostałych zadań istniało i było znaczne, co zostało stwierdzone w wyniku wielu rozmów z pracownikami urzędów, odpowiedzialnych za poszczególne zadania w projekcie, oraz w wyniku zasięgnięcia opinii ekspertów, posiadających doświadczenie w pracach projektowych o tożsamym zakresie. Dodatkowo, przy definiowaniu ryzyka zostało uwzględnione każde indywidualne zgłoszenie ze strony osób bezpośrednio zaangażowanych w projekt, tzn. urzędników, dostawców itp. Dla zwiększenia obiektywizmu identyfikacji ewentualnych rodzajów ryzyka, posłużono się również takimi technikami, jak lista kontrolna (checklist), burza mózgów, przy czym te propozycje potencjalnego ryzyka zostały poddane ocenie eksperta.

Podczas wywiadów i indywidualnych zgłoszeń osoby identyfikujące ryzyko były również proszone o podanie jego wartości, która następnie była poddana ocenie eksperta – zatwierdzana lub modyfikowana. Przy identyfikacji ryzyka należało również powiązać je z jednym lub więcej niż jednym kryterium spośród kryteriów wymienionych w rozdziale 2.

Z przeprowadzonych wywiadów wynika, że najbardziej narażonym na ryzyko opóźnienia terminu zakończenia realizacji są zadania związane ze specyfikacją wymagań oraz z integracją systemów wewnętrznych zarządzania zadaniami jednostek z platforma teleinformatyczną.

W pierwszym przypadku problem polega na formalnej formie komunikacji, wymuszonej przez Ustawę z dnia 29 stycznia 2004 r. – Prawo zamówień publicznych [15]. Ma to negatywny wpływ na efektywność komunikacji, przez co może powodować opóźnienia w wypracowaniu odpowiedniej specyfikacji wymagań, jak i również mieć negatywne konsekwencje dla rzetelności przygotowywanej dokumentacji, skutkujące w dalszej realizacji projektu.

Co się tyczy drugiego ryzykownego obszaru, problem polega na tym, iż rzeczą trudną, a wręcz niemożliwą jest dokonanie trafnej estymacji czasu potrzebnego na dokonanie integracji istniejącej infrastruktury z nowo wprowadzanym rozwiązaniem. Problemy związane z tym etapem realizacji projektu mają zróżnicowany charakter – począwszy od problemów związanych z kompatybilnością, po te, których przyczyną jest czynnik ludzki.

Potencjalne zmiany czasu realizacji poszczególnych zadań mogą sprawić, iż w trakcie realizacji projektu inne ścieżki staną się ścieżkami krytycznymi i będą decydowały o rzeczywistym czasie zakończenia projektu.

W tabeli 3 podano uzyskane wyniki oceny ryzyka, wraz z przyporządkowaniem ich od kryteriów a-e, wymienionych w rozdziale 2.

Tabela 3

Identyfikacja i kwantyfikacja ryzyka dla realizacji projektu „Rozbudowa infrastruktury teleinformatycznej na obszarze powiatu wrocławskiego i 7 gmin oraz wprowadzenie i zwiększenie dostępności elektronicznych usług dla mieszkańców i podmiotów gospodarczych regionu powiatu i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka”

Numer zadania	Definicja ryzyka	Wartość 1-5 kryterium
1	2	3
2	Brak porozumienia w kwestii podziału odpowiedzialności w projekcie	2,a)

cd. tabeli 3

1	2	3
3	Problemy komunikacyjne – trudność w uzyskaniu koniecznych informacji	3,e)
4	Opóźnienie akceptacji wniosku na szczeblu krajowym	2,e)
	Opóźnienie ostatecznej decyzji o przyznaniu dofinansowania	2,e)
	Opóźnienie w podpisaniu umowy o dofinansowanie	2,e)
5	Problemy kadrowe na poziomie każdego z partnerów projektu – brak odpowiedniej osoby do oddelegowania	2,d)
6	Brak kandydata o wymaganym profilu	2,d)
7	<i>Nieuwzględnienie wszystkich wymaganych charakterystyk</i>	4,a)
	Problemy komunikacyjne – formalizacja komunikacji – <i>Ustawa o zamówieniach publicznych</i>	4a),e)
9	Niewywiązanie się wykonawców z umowy	3,e)
	Konieczność zwiększenia zakresu prac	2,e)
10	Opóźnienia w procesie zaopatrzenia	2,e)
11	Nieodpowiedni (niezgodny z rzeczywistym zapotrzebowaniem) podział prac	2,a)
12	Problemy z kompatybilnością nowego sprzętu z istniejącą infrastrukturą, nie uwzględnione w dokumentacji technicznej	4,d)
14	Trudności w znalezieniu produktu dopasowanego do opisu zawartego w specyfikacji wymagań	2,e)
15	Nieefektywna integracja/konfiguracja systemu – niezgodna z rzeczywistymi procedurami urzędu	4,d)
16	Trudności związane ze integrowaniem rozwiązania informatyczne wielu producentów konfiguracją i integrację nowych elementów w istniejącej infrastrukturze – np. może się zdarzyć, że stosowane rozwiązania informatyczne, nie są wspierane przez producentów	5,c)
17	Braki informacyjne w istniejących bazach danych	3,a)
	Trudności z uzyskaniem spójności i integralności bazy danych	3, c)
18	Problem z wypracowaniem spójnej a zarazem funkcjonalnej wersji dokumentu	2, c)
19	Trudności z opracowaniem Polityki certyfikacji	2, c)
20	Umiejętności personelu – wymagania formalne i techniczne BIP	2, a)
	Braki w informacjach mających być opublikowanymi w ramach BIP	3,a)
21	Nieefektywna z punktu widzenia klienta integracja/konfiguracja systemu	4,c)
22	Umiejętności personelu	2,a)
23	Opór personelu przed zmianami	3,a)
24	Dostępność koniecznych informacji, dokumentacji	3,a)
	Dostępność personelu	2,b)
	Problemy komunikacyjne	3,a)

cd. tabeli 3

1	2	3
25	Opóźnienie w dostarczeniu usługi – niewielka liczba oferentów	2,e)
26	Trudności z uzyskaniem wymaganej funkcjonalności przy przestrzeganiu wymogów przyjętej Polityki Bezpieczeństwa	2,e)
27	Słaba koordynacja zasobów ludzkich	2,a),e)
	Wady ukryte produktów projektu	3,c)
28	Nieprawidłowa konfiguracja technologiczna systemu	3,c)
	Braki sprzętowe	2,e)
29	Opór personelu przed zmianami – konieczność dodatkowych szkoleń	2,a),c)
	Umiejętność personelu	3,a)
31	Problem z wypracowaniem jednolitego, ergonomicznego z punktu widzenia użytkownika interfejsu	3,c)
33	Problem z infrastrukturą telekomunikacyjną	2,e)
	Trudności w znalezieniu personelu do obsługi CT	2,e),c)
34	Umiejętność personelu	3,a),c)
35	Umiejętność personelu (sołectwa)	4 a),c)
36	Niska skuteczność prowadzonych działań – konieczność przedłużenia	2,e)
38	Opóźnienia w pracach nad wdrożeniem systemu związane z działaniami audytowymi	2,e)
40	Słaba koordynacja zasobów ludzkich	2,a)

W trakcie przyporządkowywania ryzyka do kryteriów okazało się, iż kryteria oceny stopnia krytyczności ścieżki zaproponowane w literaturze wydają się zbyt mało szczegółowe i zasugerowano, by popracować nad ich uszczegółowieniem. Wydaje się to jak najbardziej realne z praktycznego punktu widzenia, gdyż w projekcie występuje tylko dziewięć ścieżek. Proponuje się zatem wprowadzenie funkcji właścicieli poszczególnych ścieżek, którzy kontrolowaliby ryzyko wydłużenia „swojej” ścieżki ponad długość aktualnej ścieżki krytycznej. Musieliby w tym celu zapoznać się, w jakich aspektach projektu tkwią zagrożenia dla „ich” zadań, a także komunikować się z właścicielami ścieżek, które z ich ścieżkami mają część wspólną. Jednak prosta struktura sieci sprawi, iż komunikacja nie będzie zbyt rozbudowana. Wprawdzie musieliby się ze sobą komunikować niemal wszyscy właściciele ścieżek, ale komunikacja ta dotyczyłaby głównie zadań będących częścią wspólną prawie wszystkich ścieżek: zadań 2,6,5,7,3,4,21,28,29,27,40. Zadania te są aktualnie krytyczne, więc fakt, iż byłyby od początku „pilnowane” przez kilku właścicieli ścieżek, zmniejszy prawdopodobieństwo wydłużenia aktualnych ścieżek krytycznych. Jednocześnie właściciele ścieżek kontrolowaliby powiązania między zadaniami – i nie chodzi tu o powiązania uwidocznione na sieci projektu, a raczej

o zależności tam niewidoczne. Na przykład, jeśli w zadaniu 2 zmateriałizuje się ryzyko związane z umiejętnościami personelu, właściciele wszystkich ścieżek będą musieli zastanowić się, w jakim stopniu umiejętności personelu – tego konkretnego personelu z zadania 2 – będą mogły wpłynąć na zadania na „ich” ścieżkach. Właściciele ścieżek będą również mogli kontrolować, czy ich ścieżka nie okaże się połączona z innymi ścieżkami poprzez inne jeszcze powiązania, niezidentyfikowane na rys. 1. Tego typu analizy, obejmujące nie-sformalizowane powiązania, zależności oraz zjawiska pojawiające się w trakcie realizacji, mogą być z powodzeniem przeprowadzone w projekcie o takich rozmiarach i stopniu złożoności, jak analizowany tutaj projekt. Podział ryzyka według różnych kryteriów może też być pomocą: właściciele ścieżek mogą się zwracać do różnych ekspertów. Różne kryteria stanowią również podpowiedź, zwracającą uwagę właścicieli ścieżek na różne typy ryzyka, często powiązane z tymi, które wystąpiły na innych ścieżkach. Właściciele ścieżek mogą również korzystać z informacji o możliwościach przyspieszenia realizacji zadań i dzielić te informacje z innymi właścicielami ścieżek.

Na przykład ścieżka IV ma możliwości skrócenia swojego czasu trwania, dzięki dostępności dodatkowych zasobów (kryterium b)), a ścieżka IX takich możliwości nie ma. Można też zauważyć, że na ścieżkach VIII i IX nie występuje ryzyko powtarzania jakiegoś zadania, podczas gdy na pozostałych ścieżkach takie ryzyko występuje.

W przypadku zadań należących do kilku ścieżek danemu zadaniu przygląda się kilku właścicieli ścieżek, dzięki czemu mogą oni zauważyć więcej zagrożeń, ale i więcej szans z nim związanych. Powinni oni przekazywać sobie wzajemnie informacje dotyczące wspólnych zadań, a następnie każdy z nich powinien ocenić wagę i znaczenia danej informacji dla ryzyka opóźnienia „jego” ścieżki. Wymiana informacji między właścicielami ścieżek pomoże także w decyzjach o ewentualnym przerzuceniu zasobów ze ścieżek krótkich do ścieżek za długich.

Podsumowanie

W niniejszym opracowaniu zaproponowano zarządzanie ryzykiem opóźnienia projektu poprzez wprowadzenie funkcji „właścicieli ścieżek”. Właściciel ścieżki jest odpowiedzialny za jedną ze ścieżek w sieci projektowej – szacuje jej planowaną długość oraz w różnych kategoriach, ocenia ryzyko przedłużenia ścieżki w sposób zagrażający pomyślnemu zakończeniu projektu lub szanse skrócenia ścieżki tak, by zwiększyć prawdopodobieństwo zakończenia projektu w terminie.

Podejście takie nie byłoby możliwe do wdrożenia w ogólnym przypadku, gdyż sieci projektowe mogą mieć bardzo dużo ścieżek. Niemniej jednak badania pokazują, iż wiele projektów w praktyce ma sieć o prostej strukturze. Jest tak np. w wielu projektach realizowanych przez jednostki samorządowe. W przypadku tego typu sieci projektowych możliwe jest inne zarządzanie ryzykiem opóźnienia projektu – dokładniejsze, w większym stopniu oparte na komunikacji, bezpośredniej obserwacji i dzieleniu się nieformalnymi informacjami, dopasowane do potrzeb konkretnego projektu. Nie trzeba przy tym stosować złożonych modeli matematycznych, niezrozumiałych dla praktyków, a w dodatku opartych często na nieprawdziwych, a ukrytych założeniach.

Zaproponowana metoda została wstępnie przetestowana w starostwie powiatowym we Wrocławiu. Niemniej jednak nastąpiło to w zasadzie post factum, w końcowej fazie realizacji projektu. Dlatego konieczne są dalsze badania, pozwalające zweryfikować metodę w praktyce i zmodyfikować ją tak, by spełniała oczekiwania kierowników tych projektów, których sieć jest na tyle nieskomplikowana, by można było zastosować proponowaną metodę.

Zaproponowane podejście, wprowadzające dodatkowe, i to nieraz dość liczne osoby „funkcyjne” do struktury organizacyjnej projektu, a także powodujące, że niektóre zadania będą kontrolowane przez kilku właścicieli ścieżek, wymaga sprecyzowania zależności kompetencyjnych i organizacyjnych między kierownikiem projektu i właścicielami ścieżek, a także między poszczególnymi właścicielami ścieżek. Generalnie właściciele ścieżek mają włączyć się w ogólną filozofię zarządzania ryzykiem projektu, polegającą na aktywnym uczestnictwie w tym procesie możliwie wszystkich udziałowców projektu i zgłaszaniu przez nich do kierownika projektu wszelkich uwag i propozycji w tym zakresie. To kierownik projektu w ostatecznym rozrachunku zadecyduje, jakie, kiedy i czy w ogóle jakieś kroki mają być podjęte. Natomiast rola właścicieli ścieżek będzie w zarządzaniu ryzykiem opóźnienia projektu bardzo ważna, bo o opóźnieniu projektu decyduje zawsze jakaś ścieżka, która się okazuje dłuższa niż to było planowane. Dlatego tak istotne jest, by były wyznaczone osoby, które przyglądałyby się projektowi przez pryzmat ścieżek. Osoby te zbierałyby wszelkie informacje dotyczące „swoich” ścieżek, odpowiednio by je opracowywały i przekazywały kierownikowi projektu. Dzięki temu ścieżki wydłużające się lub grożące wydłużeniem zostaną wcześniej wykryte niż w przypadku metod zarządzania ryzykiem nieuwzględniających ścieżek i ich aktualnego (potencjalnie często się zmieniającego) stopnia krytyczności.

Literatura

1. Chanas S., Kamburowski J. (1981). The Use of Fuzzy Variables in PERT. *Fuzzy Sets and Systems*, 5(1). s.11-19.
2. Chanas S., Zieliński P. (2002). The Computational Complexity of the Criticality Problems in a Network with Interval Activity Times. *European Journal of Operational Research*, 136(3), s. 541-550.
3. Chanas S., Zieliński P. (2003). On the Hardness of Evaluating Criticality of Activities in a Planar Network with Duration Intervals. *Operations Research Letters*, 31(1). s. 53-59.
4. Chen S.P. (2007). Analysis of Critical Paths in a Project Network with Fuzzy Activity Times. *European Journal of Operational Research*, 183(1). s. 442-459.
5. Goldratt E. (1997). *Critical Chain*. North River Press, Great Barrington.
6. Grzeszczyszyn M. (2009). *Studium Wykonalności Rozbudowa infrastruktury teleinformatycznej na obszarze powiatu wrocławskiego i 7 gmin oraz wprowadzenie i zwiększenie dostępności elektronicznych usług dla mieszkańców i podmiotów gospodarczych regionu powiatu i gmin: Czernica, Jordanów Śląski, Kąty Wrocławskie, Kobierzyce, Mietków, Sobótka, Żórawina*. Wrocław.
7. Kamburowski J. (1992). Bounding the Distribution of Project Duration in PERT Networks. *Operations Research Letters*, 12(1). s. 17-22.
8. Kuchta D. (2002). *Miękka matematyka w zarządzaniu*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
9. Kuchta D., Kobylański P. (2007). A Note on the Paper by M.A. Al-Fawzan and M. Haouari About a Bi-objective Problem for Robust Resource-constrained Project Scheduling. *International Journal of Production Economics*, 107(2). s. 496-501.
10. Kuchta D., Ptaszyńska E. (2010). Zarządzanie ryzykiem poprzez uczenie się w projektach europejskich. W: *Modelowanie preferencji a ryzyko '10*. Red. T. Trzaskalik. AE, Katowice, s. 221-230.
11. Skorupka D. (2007). *Metoda identyfikacji i oceny ryzyka realizacji przedsięwzięć budowlanych*. WAT, Warszawa.
12. Skorupka D. (2004). Neural Networks in Risk Management of a Project. 2004 AACE International Transaction, (CSC.1.51–CSC.1.57). The Association for the Advancement of Cost Engineering, USA, Washington.
13. Valadares Tavares L., Auntes Ferreira J., Coelho J.S. (1999). The Risk of Delay of a Project in Terms of the Morphology of its Network. *European Journal of Operations Research*, 119, s. 510-537.
14. Zammori F.A., Braglia M., Frosolini M. (2009). A Fuzzy Multi-criteria Approach for Critical Path Definition. *International Journal of Project Management*, 27(3), s. 278-291.

-
15. Ustawa z dnia 29 stycznia 2004 r. – Prawo zamówień publicznych. Dz.U. 2010, nr 113, poz. 759, nr 161, poz. 1078 i nr 182, poz. 1228 oraz z 2011, nr 5, poz. 13 i nr 28, poz. 143.

MANAGEMENT OF DELAY RISK IN PROJECTS OF A SIMPLE NETWORK STRUCTURE, REALISED BY SELF GOVERNMENT UNITS

Summary

This paper focuses on the identification of the critical path in noncomplex projects networks. The main objective of this paper is to present the idea of a relatively simple but effective risk analysis of project delay for this group of projects. The proposed risk management method of project duration overrunning introduces the “path’s owners” function. This approach is based on the assumption that each of the paths in the project network is monitored by its owner who estimates its planned length and, in various categories, assesses the risk of extending the path in a way that threatens the successful project performance as well as the opportunity to shorten the path to increase probability of the on-time project completion. In order to visualize the implementation and the potential of the proposed approach, the article provides an example concerning the noncomplex networks project, performed by a local government unit in Wrocław.

Krzysztof S. Targiel

Uniwersytet Ekonomiczny w Katowicach

METODA DYNAMICZNEGO HARMONOGRAMOWANIA PORTFELI PROJEKTÓW PRZY OGRANICZONYCH ZASOBACH*

Wprowadzenie

Nowe tendencje w zarządzaniu projektami, przejawiające się w tworzeniu biur zarządzania projektami (PMO – Project Management Office), skupiających zarządzanie całym portfelem projektów danej instytucji, rodzą nowe problemy w harmonogramowaniu projektów. Na problem harmonogramowania portfela projektów można spojrzeć jak na harmonogramowanie pojedynczego projektu, w którym istnieje wiele rodzajów aktywności powiązanych w niespójne grafy czynności, połączonych jednak koniecznością korzystania ze wspólnych ograniczonych zasobów.

W tym świetle problemem staje się dobór nowych projektów do portfela oraz co za tym idzie modyfikacje istniejących harmonogramów. W przedstawionej perspektywie na problem ten można spojrzeć jak na rozważany już w literaturze problem dynamicznego harmonogramowania, polegający na modyfikacjach w sieciach czynności pojedynczych projektów. W pracy przedstawiono koncepcję dynamicznego harmonogramowania portfeli projektów przy ograniczonych zasobach pozwalające na optymalny dobór nowego projektu do istniejącego portfela ze względu na ograniczenia w dostępności do zasobów.

1. Biura projektowe

Biura Projektowe (Project Management Office – PMO) to za definicją PMBoK [15]: „wyodrębnione jednostki organizacyjne do celów zcentralizowania i koordynacji zarządzania projektami”. Ich specyfiką jest ciągła do-

* Praca naukowa finansowana ze środków na naukę w latach 2010-2013 jako projekt badawczy nr NN 111 267138.

stępnosc w oferowaniu wsparcia w zakresie zarządzania projektami w formie szkoleń, oprogramowania itp. Biura projektowe tworzą w swej działalności portfele projektów.

Jak podkreśla PMBoK, jedną z kluczowych własności PMO jest dzielenie i koordynacja zasobów pomiędzy wszystkimi projektami administrowanymi przez biuro [15]. Tworząc nowy portfel, poprzez dodawanie nowych projektów, należy tak utworzyć harmonogram, by wszystkie zasoby były optymalnie wykorzystane, a jednocześnie by projekt miał dostęp do właściwych zasobów. Zadanie to można sprowadzić do problemu harmonogramowania przy ograniczonych zasobach.

2. Harmonogramowanie przy ograniczonych zasobach

Problem harmonogramowania przy ograniczonych zasobach oznaczany jako RCPSP (Resource Constrained Project Scheduling Problem) jest szeroko rozważany w literaturze. W problemie tym stosowana jest klasyczna notacja, którą przyjmujemy za pracę Bruckera [4]:

- $V = \{0, 1, \dots, n, n+1\}$ – zbiór czynności (zadań), czynność oznaczona numerem 0 to początek projektu a $n+1$ koniec projektu, przy n czynnościach do wykonania,
- AON – standardowo czynności w pokazywane są w węzłach (stosujemy przedstawienie Activity On Node),
- $G = (V, E)$ – to graf przedstawiający sieć czynności.

W rozważanym przypadku jest on niespójny.

Przyjmując, iż czynność i poprzedza czynność j ($i \rightarrow j$) definiowane są następujące zbiory

- $Pred(j)$ – zbiór bezpośrednich poprzedników czynności j ,
- $Succ(j)$ – zbiór bezpośrednich następników czynności j ,
- p_j – czas wykonania czynności j .

Grupa oznaczeń stosowana jest do zapisu zasobów:

$\mathbb{R}^p$ – zbiór zasobów odnawialnych na przykład pracy,

$\mathbb{R}^v$ – zbiór zasobów nieodnawialnych na przykład materiałów,

$r_{j,k}^p$ – zapotrzebowanie na odnawialny zasób k przez czynność j ,

R_k^p – dostępność odnawialnego zasobu k ,

$r_{j,k}^v$ – zapotrzebowanie na nieodnawialny zasób k przez czynność j ,

R_k^v – dostępność nieodnawialnego zasobu k .

Oznaczając odpowiednio:

S_j $j = 0, 1, \dots, n, n+1$ – moment rozpoczęcia czynności j ,

C_j $j = 0, 1, \dots, n, n+1$ – moment zakończenia czynności j .

Problem sprowadza się do znalezienia harmonogramu w postaci wektora

$$S = (S_0, \dots, S_j, \dots, S_{n+1})$$

Obliczany jest także wektor zakończeń czynności (zadań)

$$C = (C_0, \dots, C_j, \dots, C_{n+1})$$

Wprowadzane są także oznaczenia dla opóźnień czynności:

d_{ij}^{min} – minimalne opóźnienie pomiędzy czynnościami i, j ,

d_{ij}^{max} – maksymalne opóźnienie pomiędzy czynnościami i, j .

Dla środowiska wieloprojektowego wprowadzimy dodatkowe oznaczenia

$$S^q = (S_0^q, \dots, S_j^q, \dots, S_{n+1}^q)$$

będzie harmonogramem dla q -tego projektu.

Dla projektu q zdefiniujemy także zbiór trwających czynności w czasie t

$$A(S^q, t) = \{j \in V \mid S_j^q \leq t \leq S_j^q + p_j\}$$

oraz wielkość zaangażowanego k -tego zasobu w q -tym projekcie

$$r_k(S^q, t) = \sum_{j \in A(S^q, t)} r_{jk}$$

W literaturze stosowana jest standardowa notacja zapisu zadań podana tutaj za pracą Bruckera [4]

$$\alpha|\beta|\gamma$$

gdzie:

α – to typ otoczenia zasobowego,

β – to charakterystyka czynności,

γ – to typ funkcji celu.

Problemy statyczne to takie, w których należy utworzyć harmonogram przy znanych parametrach. Rozważane są także problemy modyfikacji istniejących harmonogramów poprzez zmiany w liczbie wykonywanych czynności. Te problemy nazywane są harmonogramowaniem dynamicznym [5]. Były one rozważane także przez Artigues i innych [2] oraz Hao i innych [7].

Do rozwiązania problemu najczęściej stosowane są metody ewolucyjne, takie jak systemy mrówkowe [11], roje cząsteczek [9], przeszukiwanie z tabu i symulowane wyżarzanie [12], lokalne przeszukiwanie, algorytmy genetyczne

i ewolucyjne. Wykorzystywane są także specyficzne heurystyki, jak w pracach Neumanna i Zimmermanna [13] czy Tormosa i Lova [16]. Pomimo iż stanowią one znaczącą część wykorzystywanych metod, stosowane są także metody optymalizacji. Ayala i Artigues wykorzystują metody programowania liniowego [3].

3. Model matematyczny

Proponowany model jest modyfikacją uwzględniającą wieloprojektowość oraz wykorzystanie zasobów nieodnawialnych, modelu przedstawionego w pracy Artigues [2]. Jest to model przepływów międzyczynnościowych. W pracy proponuje, aby wykorzystać go do zaplanowania nowego projektu.

Zdefiniowano binarne zmienne decyzyjne

$$x_{ij}^q \in \{0,1\}, \quad \forall i \in V^q - \{n+1\}, \forall j \in V^q - \{0\}, \forall q \cup q+1$$

opisujące następstwo pomiędzy czynnością i -tą i j -tą dla q projektów już znajdujących się w portfelu oraz $q+1$ projektu dodawanego do niego.

Wprowadzona dodatkowa zmienna

$$f_{ijk}^q \in N \quad \forall i \in V^q - \{n+1\}, \forall j \in V^q - \{0\}, \forall q, \forall k$$

opisuje konieczny przepływ k -tego zasobu pomiędzy i -tą oraz j -tą operacją w q -tym projekcie.

Funkcja celu w postaci

$$\min \left(\sum_{t=0}^{C_{\max}} [R_k^o - \sum_{q \cup q+1} r_k(S^q, t)]^+ \right)$$

minimalizuje niewykorzystanie zasoby.

Pierwsze ograniczenie to koincydencje pomiędzy czynnościami każdego projektu. Kolejne ograniczenie to zależności czasowe. Trzecie ograniczenie to formalny zapis implikacji – jeśli nie ma połączenia pomiędzy czynnościami i oraz j , to nie może być przepływu zasobów pomiędzy tymi czynnościami. Kolejne ograniczenia to bilanse przepływów zasobów. Każda czynność musi otrzymać tyle zasobów, ile potrzebuje oraz wszystkie swoje zasoby przekazać innym czynnościom. Dotyczy to jednak tylko zasobów odnawialnych. Kolejne ograniczenia zapewniają, by były wykorzystane tylko te zasoby, które są dostępne. Tym razem dotyczy to wszystkich zasobów, zarówno odnawialnych, jak i nieodnawialnych.

$$\min(\sum_{t=0}^{C_{\max}} [R_k^{\rho} - \sum_{q \cup q+1} r_k(S^q, t)]^+)$$

s.t.

$$x_{ij}^q = 1 \quad \forall (i, j) \in E^q, \forall q \cup q+1$$

$$S_j^q - S_i^q - Mx_{ij}^q \geq p_i^q - M \quad \forall i \in V^q - \{n+1\}, \forall j \in V^q - \{0\}, \forall q \cup q+1$$

$$f_{ijk}^q - Nx_{ij}^q \leq 0 \quad \forall i \in V^q - \{n+1\}, \forall j \in V^q - \{0\}, \forall q \cup q+1, \forall k \in \mathcal{R}^{\rho}$$

$$\sum_q \sum_{j \in V^q - \{0\}} f_{ijk}^q = r_k(S^q, t) \quad \forall i \in V^q - \{n+1\}, \forall q \cup q+1, \forall t, \forall k \in \mathcal{R}^{\rho}$$

$$\sum_q \sum_{i \in V^q - \{n+1\}} f_{ijk}^q = r_k(S^q, t) \quad \forall j \in V^q - \{0\}, \forall q \cup q+1, \forall t, \forall k \in \mathcal{R}^{\rho}$$

$$\sum_q r_k(S^q, t) \leq R_k^{\rho} \quad \forall k \in \mathcal{R}^{\rho}$$

$$\sum_q r_k(S^q, t) \leq R_k^{\nu} \quad \forall k \in \mathcal{R}^{\nu}$$

$$x_{ij}^q \in \{0, 1\}, \quad \forall i \in V^q - \{n+1\}, \forall j \in V^q - \{0\}, \forall q \cup q+1$$

$$f_{ijk}^q \in N \quad \forall i \in V^q - \{n+1\}, \forall j \in V^q - \{0\}, \forall q \cup q+1, \forall k$$

Zmienne x to zmienne binarne, natomiast zmienne f to zmienne przyjmujące wartości naturalne (odpowiadające ilości przekazywanych zasobów). Obliczenia prowadzone są zarówno dla istniejących w harmonogramie q projektów, jak i dla nowo dodanego projektu ($q + 1$).

4. Przykład portfela projektów

Na portfel projektów można spojrzeć jak na jeden projekt o niespójnym grafie czynności takim jak przedstawiony na rys. 1 dla portfela rozważanego w literaturze przedmiotu problemu [6].

Tabela 1

Czynności portfela projektów

Czynność j	p_j	$Pred(j)$	Zasoby
<i>Projekt P1</i>			
P1-1	3,33		2A, C
P1-2	4		2A, B, C
P1-3	5	P1-1, P1-1	A, C
<i>Projekt P2</i>			
P2-1	5		A,B,C
P2-2	4	P2-1	2A,
P2-3	5	P2-1	A,2B,C
P2-4	4	P2-2, P2-3	2A,C

Źródło: [6].

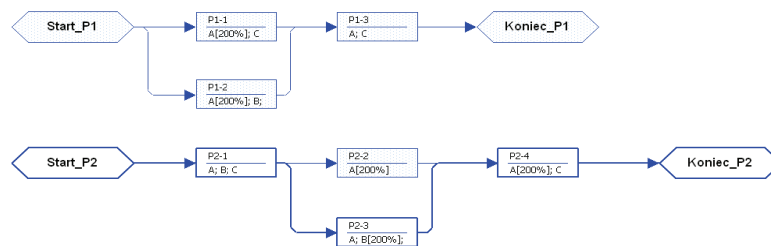
Dostępność trzech używanych zasobów (A, B, C) jest pokazana w tabeli 2.

Tabela 2

Dostępność zasobów

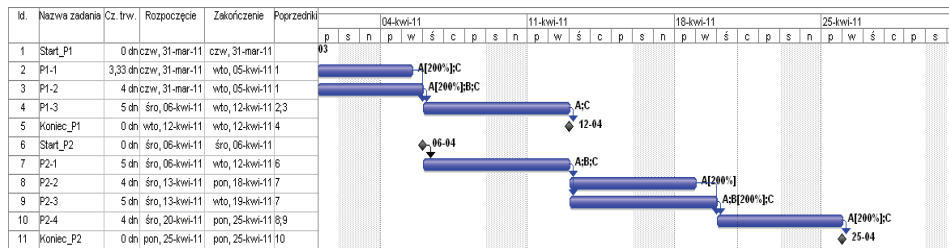
Zasób	A	B	C
Dostępność	3	2	2

Źródło: [6].



Rys. 1. Graf czynności portfela projektów

Na rys. 2. przedstawiono zapisane w programie MS Project dwa projekty z rozważanego problemu.



Rys. 2. Portfel projektów w MS Project

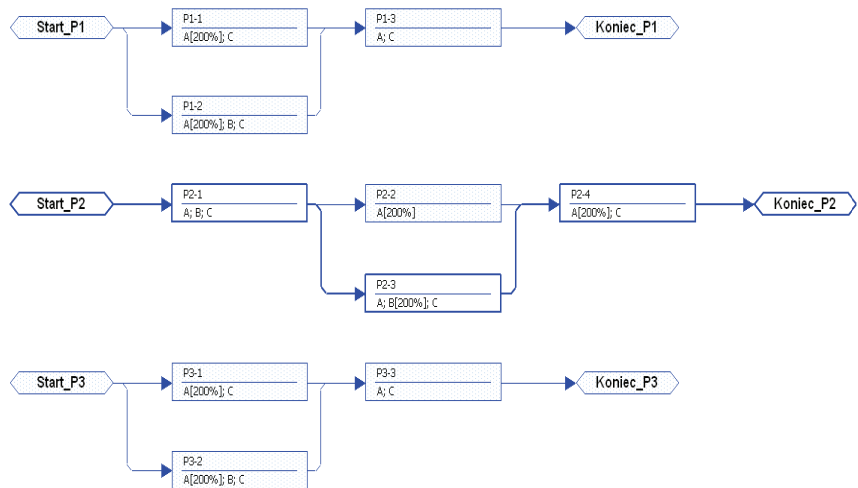
Tak zdefiniowany problem posiada zbilansowane zasoby pozwalające na realizację przedstawionych projektów w założonych czasach.

Uruchomienie przez PMO nowego projektu, powoduje konieczność zaplanowania nowych czynności, które będą korzystały z tych samych zasobów. Rozważany portfel projektów uzupełniono o projekt P3, na który składają się czynności P3-1, P3-2 oraz P3-3. Graf czynności dla tak uzupełnionego portfela przedstawiono na rys. 3.

Tabela 3

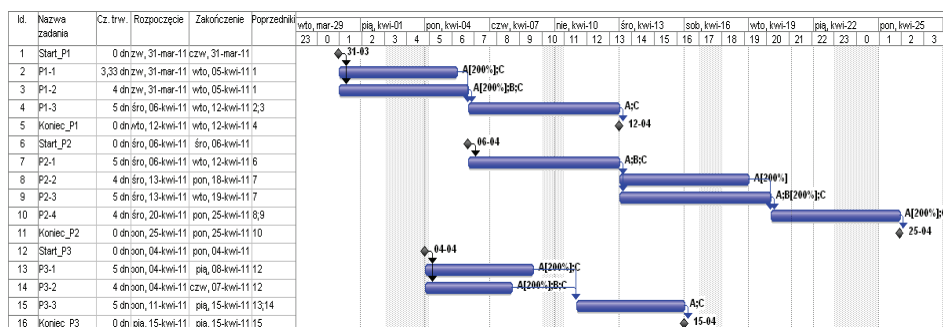
Czynności dodatkowego projektu

Czynność j	p_i	$Pred(j)$	Zasoby
Projekt P3			
P3-1	5		2A, C
P3-2	4		2A, B, C
P3-3	5	P2-1, P2-1	A, C



Rys. 3. Graf czynności uzupełnionego portfela projektów

Dodanie do portfela tego nowego projektu sprawia, iż nie posiadamy odpowiedniej ilości zasobów, by wszystkie projekty zrealizować w założonych czasach. Uzupełniony portfel projektów zapisanych w MS Project przedstawiono na rys. 4.



Rys. 4. Uzupełniony portfel projektów w MS Project

Należy tak rozmieścić zadania w czasie, by nie powodować konfliktu w dostępie do zasobów. Przy ponownym harmonogramowaniu zadań przyjęto założenia:

- zadanie rozpoczęte nie mogą być przesuwane,
- pierwszym kryterium jakości jest minimalizacja niewykorzystania zasobów przez wszystkie projekty w portfelu.

Dopiero drugim kryterium może być minimalizacja czasów przekroczenia przez wszystkie projekty naznaczonego terminu.

Gdyby założyć, iż czynności projektów już rozpoczętych nie mogą być przesuwane, problem staje się trywialny. Można wtedy sprowadzić problem do zadania, w którym tworzymy harmonogram przy zasobach ograniczonych (poprzez już wykonywane zadania) do dostępnych. Stąd dopuszczamy przesuwanie w czasie czynności projektów już rozpoczętych, ale tylko tych, które jeszcze się nie rozpoczęły.

Podsumowanie

W pracy przedstawiono ideę, założenia oraz wstępną wersję opracowywanego modelu matematycznego, w ramach projektu badawczego „Wspomaganie decyzji w planowaniu projektów”, modelu zmierzającego do wytworzenia oprogramowania usprawniającego tworzenie harmonogramów w biurach projektowych (PMO). Dalsze prace będą zmierzały do weryfikacji opracowanego

modelu poprzez jego implementację w zintegrowanym środowisku optymalizacji. Dzięki niej możliwe będzie prześledzenie działania proponowanego modelu optymalizacyjnego.

Literatura

1. Alcaraz J, Maroto C. (2001). A Robust Genetic Algorithm for Resource Allocation in Project Scheduling. *Annals Of Operations Research*, 102, s. 83-109.
2. Artigues C., Michelon P., Reusser S. (2003). Insertion Techniques for Static and Dynamic Resource-constrained Project Scheduling. *European Journal Of Operational Research*, 1, 149(2), s. 249-267.
3. Ayala M, Artigues C. (2010). On Integer Linear Programming Formulations for the Resource-constrained Modulo Scheduling Problem.
4. Brucker P., Drexel A., Mohring R., Neumann K., Pesch E. (1999). Resource-constrained Project Scheduling: Notation, Classification, Models, and Methods. *European Journal of Operational Research*, 1, s. 3-41.
5. Elkhyari A., Guéret C., Jussien N. (2003). Solving Dynamic Resource Constraint Project Scheduling Problems using new Constraint Programming Tools. *Practice and Theory of Automated Timetabling IV*, s. 39-59.
6. Hanh LQ., Rüppel U. (2008). Robust Process-based Multi-project Scheduling for Construction Projects in Vietnam. *E-Work and E-Business in Architecture, Engineering and Construction: ECPPM*, s. 327.
7. Hao Q., Xue Y., Wang S., Shen W. (2009). A Dynamic Scheduling Algorithm for Time- and Resource-Constrained Task Networks. *2009 IEEE International Conference On Systems, Man And Cybernetics*, s. 4402-4407.
8. Hao Q., Shen W., Xue Y., Wang S. (2010). Task Network-based Project Dynamic Scheduling and Schedule Coordination. *Advanced Engineering Informatics*, 10, 24(4), s. 417-427.
9. Jarboui B., Damak N., Siarry P., Rebai A. (2008). A Combinatorial Particle Swarm Optimization for Solving Multi-mode Resource-constrained Project Scheduling Problems. *Applied Mathematics And Computation*, 15, 195(1), s. 299-308.
10. Kis T. (2005). A Branch-and-cut Algorithm for Scheduling of Projects with Variable-intensity Activities. *Mathematical Programming*, 103(3), s. 515-539.
11. Merkle D., Middendorf M., Schmeck H. (2002). Ant Colony Optimization for Resource-constrained Project Scheduling. *IEEE Transactions On Evolutionary Computation*, 6(4), s. 333-346.
12. Mika M., Waligora G., Weglarz J. (2005). Simulated Annealing and Tabu Search for Multi-mode Resource-constrained Project Scheduling with Positive Discounted Cash Flows and Different Payment Models. *European Journal Of Operational Research*, 1, 164(3), s. 639-668.

13. Neumann K., Zimmermann J. (2000). Procedures for Resource Leveling and Net Present Value Problems in Project Scheduling with General Temporal and Resource Constraints. *European Journal Of Operational Research*, 1, 127(2), s. 425-443.
14. Nonobe K., Ibaraki T. (2002). Formulation and Tabu Search Algorithm for the Resource Constrained Project Scheduling Problem. *Essays And Surveys In Metaheuristics*, 15, s. 557-588.
15. Project Management Institute. (2004). *A Guide to the Project Management Book of Knowledge (PMBOK)*. Newtown Square, PA: Project Management Institute.
16. Tormos P., Lova A. (2003). An Efficient Multi-pass Heuristic for Project Scheduling with Constrained Resources. *International Journal Of Production Research*, 41(5), s. 1071-1086.

METHOD FOR DYNAMIC SCHEDULING RESOURCE-CONSTRAINED PROJECT PORTFOLIOS

Summary

The paper presents an outline of the project portfolio scheduling method, in which the portfolio was treated as a collection of inconsistent activity graph. With this approach it was able to use the achievements of the existing solutions for scheduling resource-constrained projects. The symbolism in this work set, allowed to create optimization model by which is possible to obtain solution to the problem under consideration.

Katarzyna Tyła
Tomasz Błaszczuk

Uniwersytet Ekonomiczny w Katowicach

CZAS I KOSZT REALIZACJI PROJEKTU JAKO KRYTERIA UZUPEŁNIAJĄCE TYPOWY MODEL CAD

Wprowadzenie

Komputerowe wspomaganie projektowania (Computer Aided Design, CAD) to rozwinięta klasa oprogramowania wspomagającego elektroniczne modelowanie projektowanych obiektów. Liczne aplikacje tego typu w dużym stopniu ułatwiają pracę projektanta – konstruktora oraz poprawiają jej jakość pozwalając na wirtualne modelowanie przedmiotu projektowania w przestrzeni trójwymiarowej, weryfikację rezultatów w ujęciu technicznym oraz wizualnym a także możliwość szybkiego i taniego ich przemodelowania w przypadku ewentualnej zmiany założeń lub preferencji klienta.

Procesy zarządzania projektami (PMI, 2008), poza realizacją zadań związanych z opracowaniem zakresu projektu, obejmują również wiele działań związanych z przygotowaniem kompletnego planu projektu i jego realizacją. W dalszej części pracy zaproponowane zostaną założenia modelu rozszerzającego klasyczny trójwymiarowy model otrzymywany dzięki zastosowaniu narzędzi CAD o dodatkowe dwa wymiary – czasu realizacji projektu i jego całkowitego kosztu. Uwzględnienie powyższych parametrów spełni postulat tzw. trójkąta projektu (PMI, 2008) wykazującego zależności pomiędzy zakresem projektu (tu rozumianego jako wirtualny trójwymiarowy model) a jego kosztem i terminem zakończenia realizacji. Uzupełnienie modelu projektu o powyższe dwa parametry pozwoli również na opracowanie trójkryterialnego modelu oceny wariantów projektu na bazie istniejących modeli oceny czasowo-kosztowej* (time-cost trade-off).

* Szeroki przegląd metod rozwiązywania problemów czasowo-kosztowych przedstawiono w [3].

1. Aplikacje czterowymiarowe

Informatyczne narzędzia trójwymiarowego wspomagania projektowania (AutoCAD, Bentley Microstation, TopSolid, IntelliCAD i wiele innych), są typowymi narzędziami pracy projektantów w takich dziedzinach jak inżynieria lądowa i wodna. Równie powszechne zastosowania praktyczne znajdują programy wspomagające zarządzanie szeroko rozumianymi projektami w zakresie planowania harmonogramu i budżetu. Wymienić tutaj można programy typu Microsoft Project, Primavera Project Planner, Teamwork itp. Analizując ofertę dostępnych na rynku^{*} aplikacji komercyjnych trudno jednoznacznie stwierdzić, jaki jest pełen zakres ich funkcjonalności. Nie stwierdzono jednakże oferty programu realizującego jednocześnie funkcjonalności zarówno oprogramowania klasy CAD, jak i narzędzia wspomagającego planowanie projektów w zakresie budowy harmonogramu oraz budżetu.

Technologia czterowymiarowa (4D) stanowi połączenie funkcjonalności modelowania trójwymiarowego z harmonogramowaniem prac. Model produktu statycznie opisuje techniczne i funkcjonalne cechy rezultatów projektu. Więcej informacji o technologii realizacji i jej dynamice zawiera natomiast harmonogram projektu lub jego diagram sieciowy. Modelowanie 4D wiąże dane zarówno z harmonogramu, jak i dane przestrzenne modelu trójwymiarowego. Metody wizualizacji obiektów trójwymiarowych w środowisku czterowymiarowym mogą umożliwić analizę różnych wariantów realizacji przed rozpoczęciem prac.

Możliwość tworzenia wirtualnego prototypu przed rozpoczęciem fizycznej realizacji pozwala na zmniejszenie liczby ewentualnych niezgodności oraz przyspiesza procesy projektowania i realizacji. Czterowymiarowa symulacja pozwoliłaby na postęp w zakresie elastyczności i dynamiki konstrukcji. Zaawansowane rozważania na temat możliwości aplikacji modelu czterowymiarowego prowadzili między innymi Wang i inni [11], Chau i inni [1; 2], Ma i inni [4], Russell i inni [9] oraz Mahalingam [5].

2. Koncepcja modelu pięciowymiarowego

W związku z dużymi zakresami dzisiejszych projektów oraz stopniem skomplikowania wykorzystywanych technologii, ułatwieniem mogłoby być opracowanie aplikacji komputerowej łączącej aplikacje funkcjonalności zarówno narzędzi modelowania trójwymiarowego, analizy harmonogramu i kosztów projektu. Zaspokojenie takiej potrzeby w praktyce działalności zespołów

^{*} Badanie przeprowadzono na podstawie wyników wyszukiwania popularnej wyszukiwarki internetowej oraz zestawienie zawarte pod adresem: http://en.wikipedia.org/wiki/List_of_project_management_software

projektowych wymaga jednakże opracowania stosownego modelu projektu, pozwalającego nie tylko na wirtualizację prac, lecz również zapewnienia możliwości analizy i oceny wariantów projektowych. Prace badawcze w tym kierunku zostały podjęte już między innymi przez O'Briena [6], Popov'a i innych [8] oraz Tyłę [10].

Zastosowanie koncepcji modelowania pięciowymiarowego (5D) może zapewnić efektywne zarządzanie projektem, osiągnięte dzięki możliwości stałej wzajemnej współpracy pomiędzy uczestnikami. Modelowanie to jest szansą na rozwój metod planowania z perspektywy podwykonawcy zdolnego odpowiadać za zmiany, uwzględniając wszystkie szczegóły w ogólnym planie projektu. Wdrożenie modelu pięciowymiarowego może przyczynić się do usprawnienia podejmowania decyzji w zakresie między innymi:

- wyboru kolejności realizacji zleceń,
- optymalizacji rozmieszczenia zasobów,
- analizy niewłaściwego przydziału zaplanowania zasobów i jego kosztów,
- ustalenie progów negocjacyjnych w zakresie terminów i kosztów realizacji,
- wpływu zmian technologii na wydajność/produktywność,
- alokacji zasobów pomiędzy wieloma projektami.

Złożoność i zależność powyższych zagadnień dowodzi konieczności wspomagania decyzji menedżerskich na poziomie programu lub portfela projektów.

Perspektywa analizy wykonywanych jednocześnie projektów nie unieważnia dotychczasowych badań nad modelami 4D, a jedynie proponuje ich poszerzenie. Niezbędnie jest w pięciowymiarowych modelach zawarcie informacji o pojedynczym projekcie na temat prawdopodobieństwa wystąpienia zmian w zakresie i harmonogramie, ryzyka braku postępów, możliwości wykonania alternatywnego planu pracy oraz określenie zapotrzebowania zasobów.

Modele 5D miałyby zawierać informacje pozwalające udoskonalić podejmowanie decyzji w wielu odrębnych projektach. Ponadto modele te mogłyby poprawić dokładność harmonogramu oraz pozwolić na polepszenie możliwości negocjacji kosztów, gdy konieczna jest zmiana zakresu bądź harmonogramu projektu.

Na proces konstrukcji projektu wpływa bardzo wiele czynników. Zatem, zdecydowanym ułatwieniem byłaby możliwość uwzględnienia ich wpływu już w procesie projektowania, aby zapobiec błędom mogącym wystąpić z powodu nieuwzględnienia pewnych elementów otoczenia i ich wpływu na obliczane parametry. Wysiłek włożony w kalkulacje wyznaczenia wpływu tych czynników pozwoli zaoszczędzić czas podczas konstrukcji oraz zapobiegnie niektórym pomyłkom.

Obecnie bazuje się głównie na istniejących już rozwiązaniach, które są jedynie podobne do realizowanego projektu mianowicie często korzysta się z informacji o wcześniej realizowanych projektach opierając się w ten sposób na przykładowych rozwiązaniach. Nie do końca jest to poprawne, ponieważ takie

wyliczenia dostarczają stronniczego spojrzenia na aktualne parametry. Naukowcy dążą do znalezienia takich metod, które pozwolą na popełnianie jak najmniejszych błędów w zakresie precyzji i poprawności szacowanych danych.

Do tworzenia modelu 5D niezbędne jest wykonanie w pierwszej kolejności modelu trójwymiarowego, następnie dodając czas modelu czterowymiarowego i korzystając z wcześniej zbudowanych modeli, uwzględniając kalkulacje kosztów przechodzimy do budowy modelu pięciowymiarowego. Zgodnie z założeniem o współpracy pomiędzy uczestnikami projektu konstrukcja modelu 5D [10] przedstawia się następująco:

1. W pierwszym kroku konieczne jest stworzenie modelu obiektu w 3D.
2. Kolejnym krokiem są kalkulacje czasu trwania poszczególnych czynności oraz oszacowanie czasu wykonania całego projektu, jak również skojarzenie elementów modelu trójwymiarowego z harmonogramem w celu otrzymania graficznej symulacji realizacji konstrukcji projektu.
3. Następnie należy określić zapotrzebowanie na zasoby i dokonać kalkulacji kosztów, przedstawić specyfikację produktów oraz rachunki ilościowe.
4. Czwarty krok to rozwój możliwych rozwiązań bądź alternatywnego projektu za pomocą wielokryterialnych metod podejmowania decyzji.

Wykonanie powyższych kroków prowadzi do utworzenia teoretycznego modelu informacji 3D BIM (BIM – Building Information Model [8]). Wyrażenie tego modelu w czasie podczas całego cyklu rozwojowego oraz kalkulacje kosztów są koncepcją projektowania pięciowymiarowego. Całkowity cykl jest wykonywany przez poszczególne etapy w kontekście modelowania 5D w celu osiągnięcia najlepszego rozwiązania. Informacje dotyczące cyklu przepływają pomiędzy etapami wspierając selekcję najbardziej efektywnego wariantu. Strukturalnie, model pięciowymiarowy jest podzielony na poziomy według wzoru: pierwszy poziom – prezentowane są etapy i ich współzależności, na drugim poziomie podane są początkowe dane (i procedury), trzeci poziom prezentuje zachodzące czynności i procesy, czwarty poziom pokazują uzyskane rezultaty jak również korzyści i zalety, na piątym poziomie otrzymywane są kryteria (wskaźniki) efektywności poszczególnych etapów, które są dalej używane do wielokrotnego wyznaczania wartości konstrukcji wariantów projektu.

Główną zasadą zastosowania koncepcji modelu 5D jest automatyczne uzyskanie z rozwiniętego modelu trójwymiarowego wielkości modelowanych, opisywanych i parametryzowanych elementów oraz wyznaczenie zapotrzebowania zasobów projektowych.

Oczywiste jest, że główne elementy struktury modelu mogą mieć największy wpływ na kryteria efektywności, ponieważ selekcja najlepszych wariantów zależy głównie od precyzji wyznaczenia zapotrzebowania zasobów. W wypadku jakiegoś błędu w wyznaczaniu zapotrzebowania wariant najbardziej efektywny może być niewykonalny. Dlatego w celu określenia zapotrzebowania zasobów powinny zostać wykonane dokładne obliczenia dotyczące wszystkich wielkości związanych z projektem. Modele pięciowymiarowe mogłyby zapewnić możliwość planowania z perspektywy podwykonawcy, pozwalając na efektywne i elastyczne tworzenie harmonogramu dając możliwość dokonywania zmian.

3. Wyznaczanie najbardziej efektywnego wariantu

Modele pięciowymiarowe miałyby umożliwiać nie tylko połączenie modeli 3D z kryteriami czasu i kosztu, ale także dzięki nim możliwe byłoby stworzenie kilku różnych wariantów wykonania danego projektu, a następnie stosując metody wielokryterialne możliwy byłby wybór najbardziej efektywnego z nich.

Po szczegółowym przeanalizowaniu cech projektu stworzeniu modelu 3D, poszerzeniu go o czas i koszty można przystąpić do określenia najbardziej efektywnego wariantu realizacji projektu.

Szczegółowe opisy czynności, etapów, procesów szacowania przedstawiają się następująco [10]:

1. „Blok 1 składa się z 3 etapów: pierwszy, oparty na dostępnych rysunkach i budowie modelu 3D; następnie elementy modelu trójwymiarowego są przypisywane do elementów konstrukcji i zasobów. Ostatecznie elementy są szczegółowo opisywane i parametryzowane. W ten sposób otrzymujemy komputerowo wspomagane oszacowanie kosztów. Po zebraniu wszystkich danych przechodzimy do porównywania alternatywnych rozwiązań”.

2. „Blok 2 (5D) składa się z wcześniejszego pierwszego bloku i kolejnych etapów, wymagających wyliczenia okresu wdrażania projektu i tworzenia harmonogramu. Z otrzymanych w pierwszym bloku danych i z sekwencji technologicznej otrzymujemy harmonogram z uwzględnieniem czasu trwania poszczególnych procesów oraz zapotrzebowanie na zasoby w czasie trwania konstrukcji. W tym bloku prezentowana jest także symulacja konstrukcji wdrażania projektu (wizualizacja przyszłych procesów). Kiedy wszystkie dane z bloku zostaną zebrane można przystąpić do porównania alternatywnych rozwiązań”.

3. „Blok 3 (5D) składa się z wcześniejszego bloku 1 i 2 oraz dodatkowych etapów – kontroli i zarządzania konstrukcją procesu. Wirtualny model projektu w czasie (4D) jest wykorzystywany do kontroli i porównania teoretycznego modelu i aktualnych prac konstrukcyjnych. Po zebraniu danych z tego bloku, można przystąpić do porównania alternatywnych rozwiązań i analizy ewentualnych nieregularnych dostaw zasobów”.

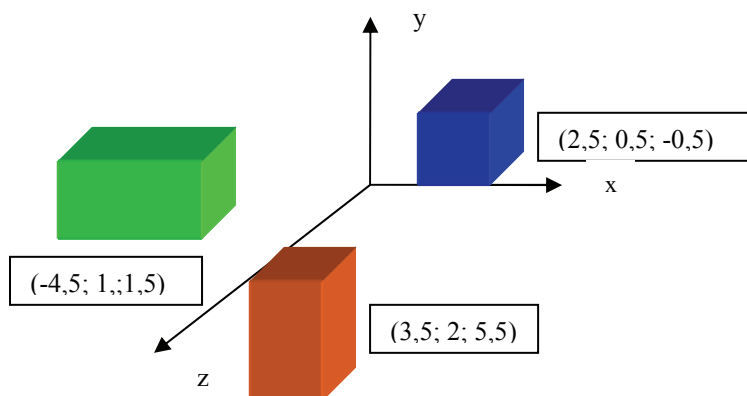
4. „Blok 4 składa się z wcześniejszych trzech bloków i kolejnych etapów – wspomagania zarządzania projektem i planowania wykorzystania projektu. W bazie danych gromadzone są wszystkie informacje dotyczące poszczególnych elementów. Gdy dane z bloku zostają zebrane, po raz kolejny przystępujemy do porównania alternatywnych rozwiązań oraz dokonujemy kalkulacji potencjalnego zapotrzebowania na zasoby w celu utrzymania ciągłości czynności”.

4. Przykład modelu pięciowymiarowego

Zastosowanie wcześniejszych rozwiązań zilustrujemy prostym przykładem obliczeniowym.

Załóżmy, że rozpatrywany problem decyzyjny dotyczy wyboru wariantu realizacji zadania polegającego na ułożenie na sobie trzech obiektów geometrycznych, których pozycja oraz przesunięcia mierzone będą miarą euklidesową w klasycznym trójwymiarowym układzie współrzędnych. Kolejność obiektów jest dowolna, natomiast warianty będą podlegały dwukryterialnej ocenie czasowo-kosztowej.

Niżej w trójwymiarowym układzie współrzędnych przedstawiono obiekty A (niebieski), B (zielony) i C (czerwony), jako współrzędne lokalizacji przyjmijmy środki ciężkości poszczególnych obiektów.



Rys. 1. Obiekty w trójwymiarowym układzie współrzędnych

Tabela 1

Odległości pomiędzy obiektami względem osi x

Obiekty	Odległość względem osi x		
	A	B	C
A	0	7	1
B	7	0	8
C	1	8	0

Tabela 2

Odległości pomiędzy obiektami względem osi y

Obiekty	Odległość względem osi y		
	A	B	C
A	0	0,5	1,5
B	0,5	0	1
C	1,5	1	0

Tabela 3

Odległości między obiektami względem osi z

Obiekty	Odległość względem osi z		
	A	B	C
A	0	2	6
B	2	0	4
C	6	4	0

Do wykonania niezbędnych obliczeń konieczne jest przyjęcie pewnych założeń dotyczących prędkości i kosztów, co zostało przedstawione w tabeli 4

Tabela 4

Koszty i prędkość przesunięć

	Prędkość (cm/s)	Koszt (zł/1g/1cm)
Przesunięcie w poziomie	10	5
Przesunięcie w pionie	5	6

Efektywny wariant realizacji zadania powinien spełniać poniższe ograniczenia

$$\text{Koszt całkowity} < 770 \text{ zł}$$

$$\text{Czas} < 2,5 \text{ s}$$

Posiadając informacje dotyczące poziomych i pionowych odległości pomiędzy obiektami należy obliczyć czas i koszty. Zakładając stałą prędkość podaną w tabeli 4 czas przesunięcia obiektu obliczamy ze wzoru

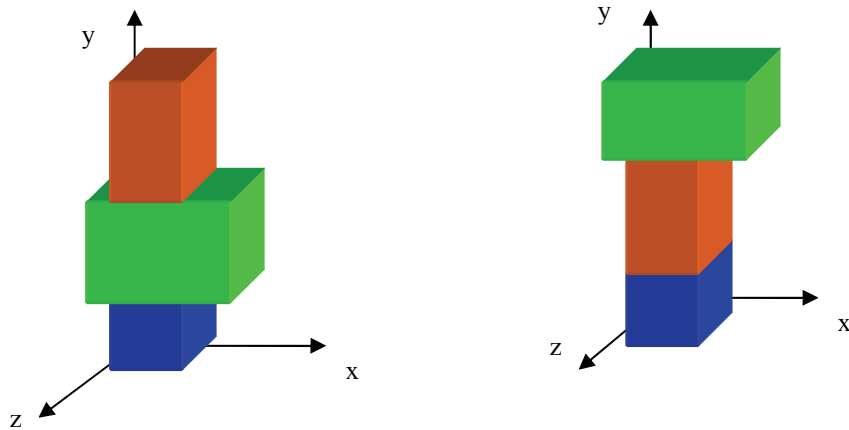
$$t = \frac{s}{v}$$

gdzie:

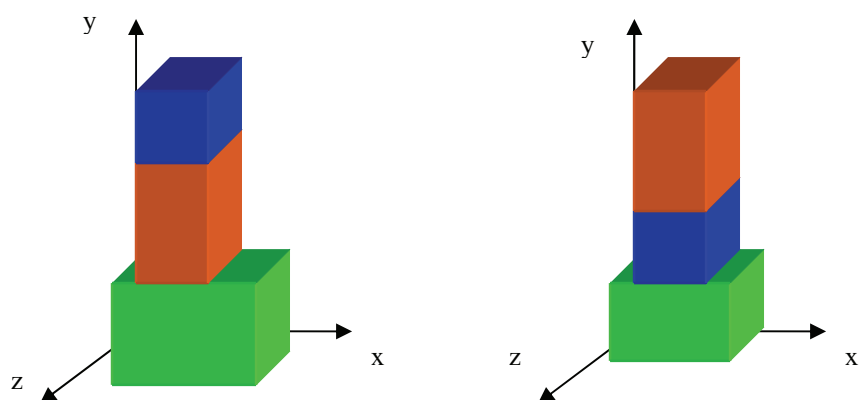
t – czas,
s – droga (przesunięcie),
v – prędkość.

Koszt łączny natomiast obliczamy sumując koszty przesunięcia obiektów względem osi x, y oraz z.

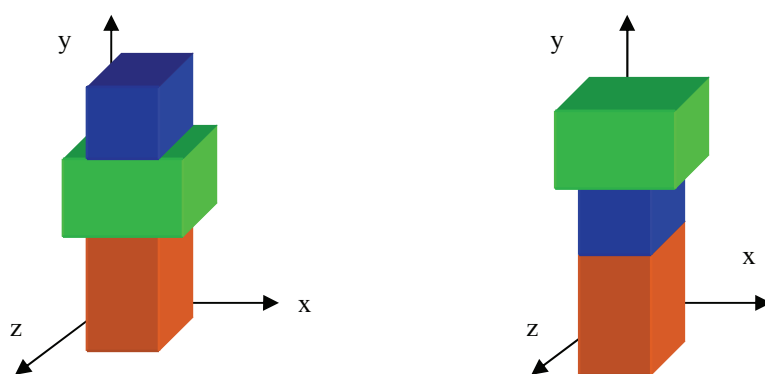
Rozważać będziemy sześć wariantów. Możliwe kombinacje oznaczmy jako: ABC, ACB (rys. 2), BCA, BAC (rys. 4), CBA, CAB (rys. 5).



Rys. 2. Ustawienie obiektów w kolejności ABC, ACB



Rys. 3. Ustawienie obiektów w kolejności BCA, BAC



Rys. 4. Ustawienie obiektów w kolejności CBA, CAB

Tabela 5

Zestawienie wyników dla wszystkich możliwych kombinacji ustawienia obiektów

Możliwe ustawienia obiektów	Przesunięcie względem osi x	Przesunięcie względem osi y	Przesunięcie względem osi z	Czas (s)	Koszt łączny
ABC	10,5	2,5	7,5	2,3	768
ACB	10,5	3	7,5	2,4	780
BAC	10,5	2,5	7,5	2,3	762
BCA	10,5	4	7,5	2,6	756
CBA	10,5	5	7,5	2,8	768
CAB	10,5	4,5	7,5	2,7	774

Z tabeli 5 wynika, że w analizowanym przypadku wyróżnić można jedynie dwa warianty dopuszczalne: ABC i BAC. Warianty ACB, BCA oraz CBA nie spełniają natomiast jednego z ograniczeń. Ze względu na identyczną wartość kryterium czasu w obydwu wariantach dopuszczalnych decydent wybierze zapewne wariant BAC, charakteryzujący się niższym kosztem realizacji i dominującym wariant ABC. W sytuacji braku oczywistej dominacji któregoś z wariantów dopuszczalnych w celu wyboru wariantu realizacji należy zastosować jedną z wielokryterialnych metod decyzyjnych, w szczególności metod stosowanych w analizie czasowo-kosztowej wariantów projektowych.

Powyższy, uproszczony model zawiera jedynie kilka niezbędnych do wykonania obliczeń informacji. Brak większej liczby informacji dotyczących kosztów, materiałów, zapasów, pracowników powoduje, że podejmowane decyzje opierane są tylko na podstawie oszacowanego czasu i podstawowych kosztów. W sytuacji, gdy do wyboru mamy tylko dwa warianty zdecydowanie prościej jest podjąć decyzję przy ograniczonych informacjach. W chwili kiedy rozważanych opcji pojawia się więcej, a rozbieżności w wynikach są wyraźnie widoczne podejmowanie decyzji staje się bardziej problematyczne i w rzeczywistości wymaga zdobycia obszerniejszych informacji dotyczących realizacji projektu.

Najkorzystniejsze byłoby w przypadku tego rodzaju projektu zastosowanie jednego z programów służących do rozwiązywania zadań optymalizacyjnych. Przykładem może być wbudowany w arkusz kalkulacyjny MsExcel program Solver lub programy AIMMS albo pakiet SAS (zapewniających możliwość tworzenia wykresów czasu wykonania poszczególnych zadań).

Dzięki dostępnej technologii obliczeniowej złożoność projektów nie stanowi problemu. Najważniejszym problemem pozostaje brak dostępnego oprogramowania dedykowanego.

Podsumowanie

Złożoność projektów wymaga zazwyczaj wykonania wielu działań, wynikających i powiązanych ze sobą przyjętą technologią realizacji w celu osiągnięcia zamierzonych rezultatów.

Dostępne na rynku aplikacje służące do komputerowego wspomaganie projektowania pozwalają na tworzenie trójwymiarowych obiektów lub harmonogramu. Są to jednak niepowiązane ze sobą programy, które dają możliwość powiązania obiektu z jego parametrami lub połączenia harmonogramu, kosztów i zasobów. Opracowanie to poświęcone zostało tematyce modelowania pięciowymiarowego, a ściślej założeniom modelu aplikacji komputerowej, umożliwiającej równoległą analizę zakresu projektu, jego budżetu i harmonogramu. Korzyści osiągnięte dzięki automatyzacji obliczeń w zastosowaniu modelu pięciowymiarowego podczas planowania projektu dotyczą w pierwszej kolejności ułatwienia zarządzania jego integracją i konfiguracją.

Modelowanie pięciowymiarowe w praktyce mogłoby w korzystny sposób wpłynąć na jednoczesne modelowanie harmonogramu i budżetu pozwalając zaoszczędzić czas oraz zminimalizować liczbę potencjalnych błędów. Również ważną zaletą tego rodzaju oprogramowania byłaby możliwość symulowania wariantów wykonania projektu i ich oceny w kontekście trzech kryteriów – zakresu, czasu i kosztów realizacji.

Można zatem stwierdzić, że niewątpliwie innowacyjne podejście w zarządzaniu projektami, jakim może stać się modelowanie 5D, mogłoby dać szansę na poprawę jakości, szybkości oraz dokładności prac projektowych.

Literatura

1. Chau K.W., Anson M., Zhang J.P. (2005). 4D Dynamic Construction Management and Visualization Software: 1. Development. *Automation in Construction*, 14, 4, 20th International Symposium on Automation and Robotics in Construction: The Future Site, s. 512-524.
2. Chau K.W., Anson M., De Saram D.D. (2005). 4D Dynamic Construction Management and Visualization Software: 2. Site Trial. *Automation in Construction*, 14, 4, 20th International Symposium on Automation and Robotics in Construction: The Future Site, s. 525-536.
3. Gasparis-Wieloch H. (2009). Wyznaczanie wariantów czasowo-kosztowych realizacji projektu. W: *Badania operacyjne w planowaniu projektów*. Red. T. Trzaskalik. AE, Katowice, s. 33-70.

4. Ma Z., Shen Q., Zhang J. (2005). Application of 4D for Dynamic Site Layout and Management of Construction Projects. *Automation in Construction*, 14, 3, International Conference for Construction Information Technology 2004, s. 369-381.
5. Mahalingam A., Kashyap R., Mahajan C. (2010). An Evaluation of Theapplicability of 4D CAD on Construction Projects. *Automation in Construction*, 19, 2, s. 148-159.
6. O'Brien W. (2000). Towards5D CAD – Dynamic Cost and Resource Planning for Specialist Contractors. *Construction Congress VI: Building Together for a Better Tomorrow in an Increasingly Complex World*, s. 1023-1028.
7. Project Management Institute. *A Guide to the Project Management Body of Knowledge*. (2008). PMI, USA.
8. Popov V., Juocevicius V., Migilinskas D., Ustinovichius L., Mikalaukas S. (2010). The Use of a Virtual Building Design and Construction Model for Developing an Effective Project Concept in 5D Environment. *Automation in Construction*, 19, s. 357-367.
9. Russell A., Staub-French S., Tran N., Wong W. (2009). Visualizing High-rise Building Construction Strategies using Linear Scheduling and 4D CAD. *Automation in Construction*, 18, 2, s. 219-236.
10. Tyła K. (2009). Modelowanie projektów w ujęciu pięciowymiarowym. W: *Zastosowanie metod ilościowych w przedsiębiorstwie*. Red. P. Miszczyński. Wydawnictwo Uniwersytetu Łódzkiego, Łódź, s. 69-74.
11. Wang H.J., Zhang J.P., Chau K.W., Anson M. (2004). 4D Dynamic Management for Construction Planning and Resource Utilization. *Automation in Construction*, 13, 5, Current IT Research and Development in the Construction Industry of China, s. 575-589.
12. http://en.wikipedia.org/wiki/List_of_project_management_software (15.04.2010).

PROJECT DURATION AND COST AS A COMPLEMENTARY CRITERIA FOR THE STANDARD 3D CAD MODEL

Summary

The aim of this paper was a compilation the proposal of five-criterial optimization model supporting project planning processess. The research has been conducted with the assumption of the validity of the linkage between standard bi-criterial time-cost trade-off model and additional three dimensional criteria followed by the designing methodology baded on Computer Aided Designing (CAD) tools. Results are illustrated by the simple numerical example.

Łukasz Tync

Independ ERP Sp. z o.o.

PODEJŚCIE OBIEKTOWE I WZORZEC OBSERWATOR JAKO WSPARCIE DLA ZARZĄDZANIA RYZYKIEM I KOMUNIKACJĄ W PROJEKCIE

Wprowadzenie

Truizmem jest twierdzenie, że współczesne organizacje o zasięgu globalnym prowadzą wiele skomplikowanych, często powiązanych ze sobą projektów. Trudno znaleźć w takim przypadku projekt, na który nie miałoby wpływu otoczenie, z innymi projektami włącznie. Środowisko projektowe staje się coraz bardziej skomplikowane, rośnie liczba uczestników, interesariuszy poszczególnych projektów, jak również powiązań pomiędzy samymi projektami. Rośnie ilość wykorzystywanych metodyk i narzędzi mających na celu zapewnienie jak najlepszej kontroli poszczególnych obszarów zarządzania projektem i wspomaganie komunikacji równocześnie gwarantując odpowiedni poziom kontroli ryzyka. Również projekty same w sobie stają się coraz bardziej skomplikowane, w wielu przypadkach realizowane przez kilka współpracujących ze sobą zespołów o różnej kulturze organizacyjnej. Taka sytuacja pozwala na określenie środowiska projektowego wielu organizacji jako heterogenicznego.

Jednocześnie z roku na rok zwiększa się tempo zmian zachodzących w organizacji i w jej otoczeniu. W trakcie trwania projektu wielokrotnie zmieniają się parametry kluczowe dla jego sukcesu. Zespół projektowy musi być zarządzany na tyle dynamicznie i elastycznie, żeby za tymi zmianami nadążyć.

Na pierwszy rzut oka można tutaj zauważyć konflikt pomiędzy rozbudowanymi procedurami i koniecznością ustrukturalizowania i ustandaryzowania wielu czynności, a – co najmniej tak samo istotną – szybkością reakcji i dostosowania się do zmieniających się realiów.

Mało tego, projekt jest często realizowany przez tzw. organizacje wirtualne, które łączą kilka przedsiębiorstw na potrzeby tylko jednego projektu, a po jego zakończeniu są rozwiązywane. Każda z firm posiada własne metodyki bądź tzw. najlepsze praktyki (best practices), z których nie chce zrezygnować.

Stawia to przed kierownikiem projektu, a co za tym idzie, również przed używanymi przez niego narzędziami coraz to nowe wyzwania – konieczność elastycznego reagowania na zmiany, przy zachowaniu niezbędnych standardów i procedur.

Z analogicznym problemem spotkali się wcześniej twórcy oprogramowania. Lawinowo rosnący poziom skomplikowania tworzonego oprogramowania, wzrastająca liczebność zespołów programistycznych, konieczność elastycznego i szybkiego reagowania na zmiany pociągnęły za sobą zmianę podejścia do tworzenia oprogramowania. Wprowadzono programowanie obiektowe (Object Oriented Programming – OOP), w którym tworzony program, zbudowany z szeregu luźno ze sobą powiązanych obiektów, miał możliwie najlepiej modelować pewien wycinek rzeczywistości. „Modele te [...] reprezentują różne aspekty obiektów, z których pochodzą ich funkcje, oraz właściwości i ich wzajemne oddziaływanie” [5]. Nowe podejście pozwoliło na tworzenie dużo bardziej rozbudowanych programów o przejrzystej strukturze, zachowujące równocześnie dużą elastyczność. Bardzo istotna tutaj okazała się „możliwość zapanowania nad chaosem, towarzyszącym nieodłącznie tworzeniu aplikacji” [5]. Podobne korzyści można uzyskać adaptując pewne elementy podejścia obiektowego do celów zarządzania projektami.

Niniejszy tekst w sposób bardzo ogólny przedstawia podstawowe założenia podejścia obiektowego w zarządzaniu projektami, a następnie skupia się na jego implementacji w obszarze zarządzania komunikacją i ryzykiem w projekcie. Zaprezentuje się – również zaadoptowany z teorii programowania – wzorzec obserwator, a następnie wskaże korzyści, które jego implementacja w dużym, heterogenicznym środowisku projektowym może przynieść. W podsumowaniu znajdzie się także propozycja dalszych kierunków rozwoju koncepcji podejścia obiektowego.

1. Przegląd literatury

Autor niniejszego artykułu, pierwszej próby przeniesienia podejścia obiektowego na grunt zarządzania projektami, a w szczególności zarządzania czasem w projekcie, podjął się w swojej pracy magisterskiej oraz w pracy zbior-

rowej pod redakcją T. Trzaskalika [6]. W ramach pracy magisterskiej stworzono także edukacyjny program komputerowy wykorzystujący zaprezentowane podejście w harmonogramowaniu projektów metodą CPM i PERT.

Niniejsze opracowanie jest pierwszą publikacją po kilkuletniej przerwie w pracach, dlatego wciąż koncepcja znajduje się we wczesnej fazie rozwoju i pełna analiza literatury nie została jeszcze wykonana. Dotychczas znaleziono nieliczne publikacje, podejmujące podobną tematykę.

Zorientowane obiektowo techniki pojawiały się głównie w zarządzaniu projektami budowlanymi, przy czym wydaje się, że obiektowość jest tutaj wykorzystywana do zamodelowania obiektu, który ma być produktem projektu budowlanego [3]. W ciekawy sposób wykorzystano również podejście obiektowe do harmonogramowania projektów powtarzalnych (repetitive projects), gdzie model obiektowy pozwolił na bardziej elastyczne harmonogramowanie, modelując równocześnie miękkie relacje pomiędzy poszczególnymi zadaniami (tzw. soft logic) [1; 2] wynikające na przykład z ograniczeń określonych zasobów.

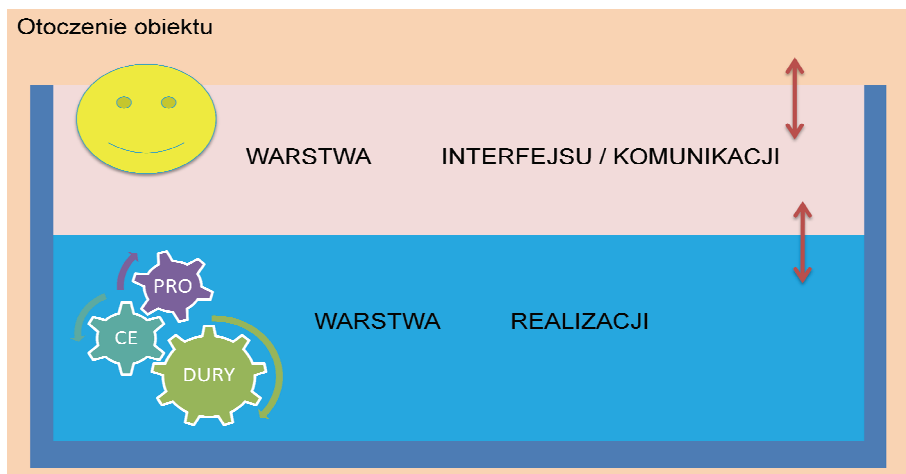
Nie znaleziono jeszcze publikacji, która postulowałaby budowę obiektowego modelu całego środowiska projektowego w taki sposób, jaki postulowany jest przez autora niniejszej koncepcji.

2. Model środowiska projektowego jako sieć luźno powiązanych ze sobą obiektów

W programowaniu obiektowym duży nacisk położono na stworzenie elastycznego modelu odwzorowującego pewien wycinek rzeczywistości. Model ten jest budowany z luźno powiązanych, komunikujących się ze sobą obiektów, które powinny mieć ściśle określony zakres odpowiedzialności oraz pewien poziom autonomii, wraz z zestawem zachowań (metod), które mogą wykonywać. Co niezwykle istotne, poszczególne obiekty mogą być składane z mniejszych obiektów (kompozycja) bądź być tworzone na wzór już istniejących, przejmując większość ich cech, jednakże możliwe jest ich rozbudowanie o nowe cechy, bądź zachowania (dziedziczenie).

Podejście Obiektowe będzie omawiane szczegółowo w kolejnych referatach, w tym miejscu zostaną tylko wprowadzone elementy niezbędne dla tematu tego opracowania.

Rysunek 1 przedstawia warstwową strukturę charakterystyczną dla każdego obiektu. W warstwie interfejsu będą się znajdować wszelkie zachowania (metody) odpowiedzialne za komunikację obiektu z otoczeniem zewnętrznym, natomiast w warstwie realizacji zaimplementowane są rzeczywiste algorytmy realizujące zakres odpowiedzialności obiektu. Poszczególne warstwy mogą być również zbudowane z innych obiektów.



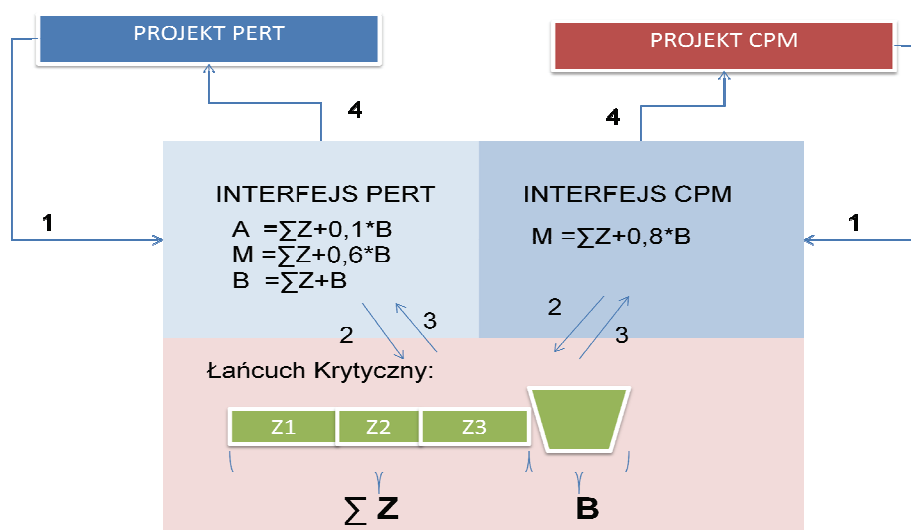
Rys. 1. Warstwowa struktura obiektu

Przykładem takiego obiektu może być nawet pojedynczy projekt, podprojekt czy też czynność. W takim przypadku w warstwie interfejsu znajdują się elementy niezbędne do komunikacji z otoczeniem danego projektu, a warstwa realizacji, zawierająca rzeczywistą implementację poszczególnych działań, będzie ukryta przed otoczeniem tego projektu.

Na rys. 2 przedstawiono schemat komunikacji podprojektu harmonogramowanego metodą łańcucha krytycznego w projekcie wykorzystującym inne metodyki. Dla uproszczenia wzięto pod uwagę jedynie czas realizacji projektu.

W tym przypadku w warstwie realizacji, znajdzie się harmonogram z metody Łańcucha krytycznego wraz z metodami lub obiektami odpowiedzialnymi za ten harmonogram, natomiast w warstwie interfejsu znajdziemy interfejsy dla harmonogramów CPM i PERT.

Obiekty odwołujące się z zewnątrz, będą używać tych interfejsów, aby uzyskać niezbędną informację, natomiast to interfejsy powinny „wiedzieć” w jaki sposób przetworzyć dane udostępniane wewnątrz obiektu, aby uzyskać oczekiwany przez obiekty zewnętrzne rezultat.



Rys. 2. Przykład – schemat komunikacji podprojektu harmonogramowanego metodą łańcucha krytycznego z projektami CPM i PERT

1. Projekt przesyła zapytanie o czasy realizacji do podprojektu (zadania Z).

2. Odpowiedni interfejs przejmuje inicjatywę. Interfejs „wie” jak skalkulować niezbędne czasy, w tym celu wysyła zapytanie o potrzebne dane do harmonogramu.

3. Niezbędne informacje są przesyłane zwrótnie do interfejsu.

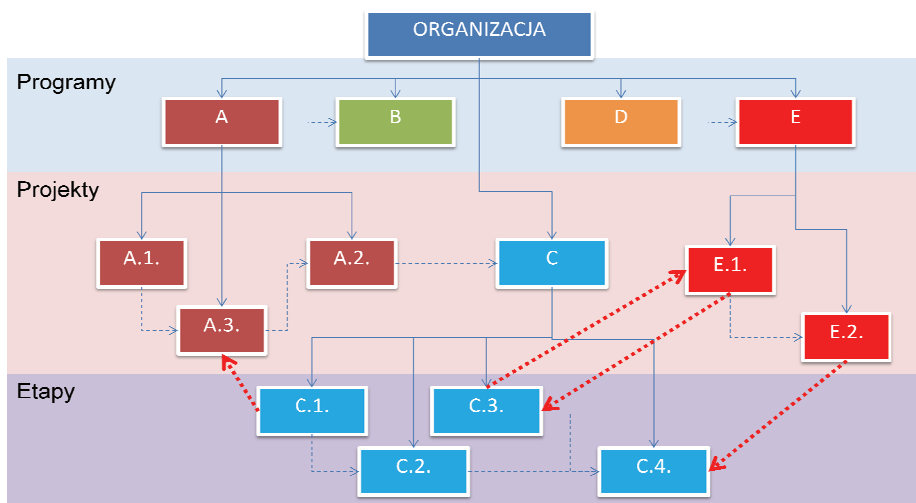
4. Interfejs kalkuluje czas realizacji zadania i przekazuje do obiektu, który przesłał zapytanie.

W tym przykładzie interfejs PERT jako czas optymistyczny podaje czas realizacji zwiększony o 10% bufora projektu ($A = \sum Z + 0,1 \cdot B$), jako czas realny czas realizacji plus 60% projektu ($M = \sum Z + 0,6 \cdot B$), a jako czas pesymistyczny czas realizacji zwiększony o 100% bufora projektu ($B = \sum Z + B$). Interfejs CPM jako czas realizacji projektu podaje czas realizacji projektu z metody łańcucha krytycznego zwiększony o 80% bufora projektu ($M = \sum Z + 0,8 \cdot B$).

3. Wzorzec obserwator – broker informacji

Odwzorowaniem dużych środowisk projektowych będzie wielopoziomowa struktura poczynszy od najważniejszych programów lub projektów, a na najbardziej szczegółowych zadaniach skończywszy. Powstała struktura będzie

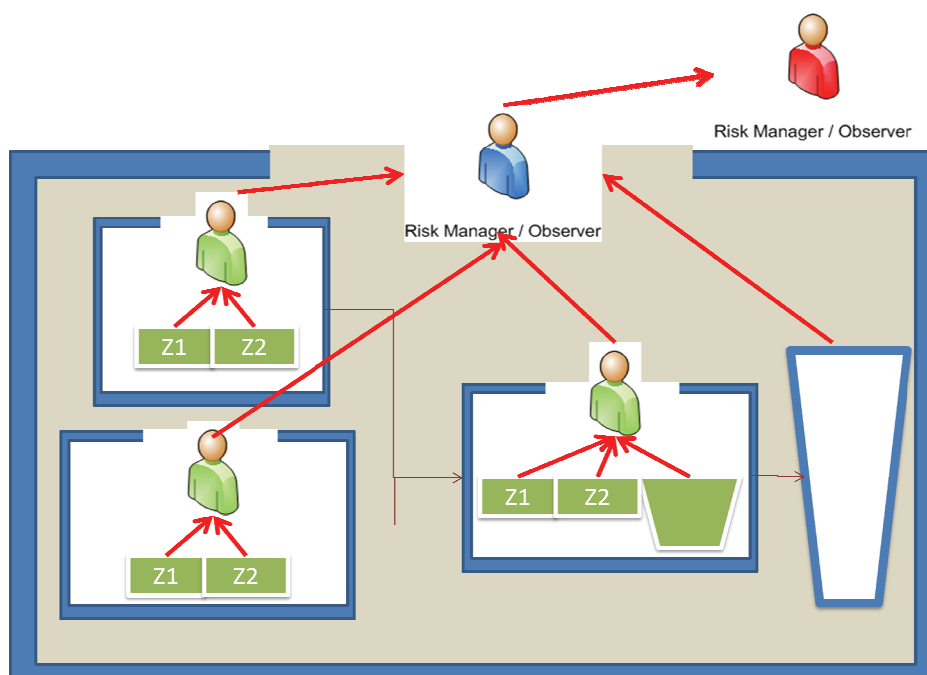
miała drzewiasty kształt, gdzie obiekty znajdujące się niżej (de facto będące elementem obiektów nadrzędnych) są uszczegółowieniem obiektów nadrzędnych. Przykładowy fragment takiej struktury zaprezentowano na rys. 3. Linie ciągłe ilustrują relację podrzędności, linie przerywane następstwo, natomiast linią kropkowaną zaznaczono inne przykładowe zależności, najczęściej są to różnego rodzaju powiązania pośrednie (wynikające np. ze współdzielenia zasobów).



Rys. 3. Struktura sieci projektów

Im bardziej rozbudowane jest środowisko projektowe danej organizacji, tym więcej różnego rodzaju powiązań pojawia się pomiędzy obiektami. Oznacza to, że poszczególne obiekty powinny informować się nawzajem o zmianie swojego stanu (kluczowych parametrów, jak np. czas trwania czy moment rozpoczęcia czynności). Im więcej jest takich powiązań, tym trudniej zapanować nad nimi i trudniej zapewnić właściwy przepływ informacji. Szczególnie trudne do kontrolowania są powiązania pośrednie, co może być istotnym czynnikiem zwiększającym ryzyko projektu.

Pewną pomocą w takich przypadkach może być – przeniesiony również z teorii informatyki – wzorzec obserwator. Zgodnie z prezentowaną koncepcją będzie to obiekt bądź – jak w przedstawionym poniżej przykładzie – sieć obiektów odpowiedzialnych za przepływ informacji o zmianie stanów poszczególnych obiektów. W obiekcie tym również mogą zostać zaimplementowane mechanizmy zarządzania ryzykiem.



Rys. 4. Wzorec Obserwator

Podstawowym zadaniem obserwatora jest informowanie o zmianie stanu obiektów, za które odpowiada (np. w ramach danego etapu projektu). Każdy z obiektów „wie”, który obiekt jest odpowiedzialnym za dany obszar obserwacji i powiadamia o zmianie swojego stanu jedynie ten obiekt. Obiekt obserwator zaś prowadzi swego rodzaju listę obiektów zainteresowanych informacjami o stanie poszczególnych obiektów obserwowanych i w momencie zmiany stanu jednego z obiektów rozsyła informację o tym do wszystkich zainteresowanych obiektów. Każdy obiekt może się na taką listę zarejestrować, bądź z niej wyrejestrować, jeśli informacje o zmianie stanu przestają być mu już potrzebne.

Dobłą – z życia wziętą – ilustracją wzorca obserwator może być tzw. newsletter dostępny na wielu stronach WWW. Każdy z odwiedzających – zainteresowany publikowanymi na stronie informacjami – może się wpisać na listę mailingową, a mechanizm newslettera będzie go informować o nowościach na danej stronie (czyli o zmianie jej stanu).

4. Obserwator – Menedżer Ryzyka

Obiekt Obserwator może być również rozbudowany o mechanizmy kontroli ryzyka, wspomagające identyfikację ryzyka wraz z systemem wczesnego ostrzegania.

Każdy obiekt może mieć własne reguły służące ocenie ryzyka na danym poziomie. Taka ocena, w ustandaryzowanej formie może być przekazywana do najbliższego obserwatora, który przekazuje dalej zagregowaną informację.

Obiekt Obserwator z kolei analizuje spływające do niego informacje i – na podstawie zestawu reguł – szacuje poziom ryzyka w obszarze, za który odpowiada oraz „podejmuje decyzje” o dalszej reakcji. Niezbędne informacje są także przekazywane na wyższe poziomy, co pozwala szacować ryzyko dla całych projektów, programów, a nawet całej organizacji.

Zakładając, że całość jest wspierana przez system informatyczny, można sobie wyobrazić, że nie musi to być jedynie prosta agregacja, ale także bardziej zaawansowana analiza. Pojawienie się kilku „żółtych świateł” może być inaczej traktowane na ścieżce krytycznej, a inaczej poza nią. Nagły „wysyp” kilku takich przypadków też może być dodatkową informacją, że z projektem dzieje się coś złego.

Dodatkową zaletą mogą być ustrukturalizowane informacje o mniej oczywistych powiązaniach pomiędzy obiektami. Jeżeli projekty/zadania (obiekty), wydawałoby się zupełnie niepowiązane, zaczynają „interesować się” sobą nawzajem, być może jest to sygnał, że istnieje jakaś zależność pomiędzy nimi (poprzez współdzielone zasoby, nie tylko ludzkie, ale również np. informatyczne). Autor nie prowadził dotychczas badań w tym kierunku, ale być może ciekawe efekty dałoby zaadoptowanie metod znanych z socjometrii. Być może zbudowanie modelu sieci takich powiązań i poddanie go analizie wskazałoby dodatkowe czynniki ryzyka.

Taka sieć obserwatorów zawierająca zestaw reguł identyfikujących czynniki ryzyka oraz wzbogacona o gromadzoną w każdym projekcie wiedzę (tzw. lessons learned), pozwoliłaby stworzyć czuły oraz coraz dokładniejszy system wczesnego ostrzegania organizacji realizującej wiele projektów.

5. „Uczące się” środowisko projektowe – baza wiedzy

Kolejne aspekty podejścia obiektowego widoczny jest głównie w przypadku implementacji na podstawie odpowiednich narzędzi informatycznych. Koncepcja zakłada budowę modelu środowiska projektowego, który będzie

wykorzystywany w wielu projektach organizacji. Elastyczność modelu pozwala traktować go jako swego rodzaju bazę wiedzy na temat zrealizowanych już projektów oraz stosowanych w organizacji metodyk. Może się to odbywać na różnych poziomach.

Tworzone jest repozytorium zawierające definicje poszczególnych typów obiektów wykorzystywanych w danym środowisku projektowym (typy te – za podejściem obiektowym – nazywane będą klasami). Rozbudowane obiekty będą zawierać w sobie zdefiniowane elementy logiki biznesowej organizacji (np. sposoby translacji danych pomiędzy metodykami czy reguły oceny ryzyka).

Innym źródłem wiedzy może być historyczny zapis wszystkich komunikatów informujących o zmianie stanu poszczególnych obiektów przesyłanych do obserwatorów. Analiza takiego zapisu historycznego za pomocą metod heurystycznych czy też pewnych narzędzi Business Intelligence może pozwolić na identyfikację mniej oczywistych, charakterystycznych dla organizacji schematów ostrzegających przed wystąpieniem określonego ryzyka.

Podsumowanie

Wyżej bardzo ogólnie zaprezentowano koncepcję podejścia obiektowego do zarządzania projektami. Koncepcja ta będzie rozszerzana w kolejnych pracach. Zaprezentowano również wzorzec Obserwator jako jeden z elementów prezentowanej koncepcji, który może mieć zastosowanie w zarządzaniu ryzykiem w trakcie realizacji projektu.

Do zalet zaprezentowanej koncepcji, można zaliczyć:

1. Ustrukturalizowanie całego środowiska projektowego.
2. Jasne reguły łączenia różnych metodyk (na różnych poziomach organizacji lub dla różnych obszarów zarządzania projektami).
3. Możliwość rozbudowy o własne narzędzia oraz przejrzyste łączenie ich z istniejącymi procedurami.
4. Poprawę komunikacji (zwłaszcza pomiędzy projektami).
5. Możliwość identyfikacji i kontroli również nieoczywistych powiązań pomiędzy projektami.
6. „Uczącą się” strukturę.
7. Możliwość integracji z otoczeniem zewnętrznym.
8. Możliwość automatyzacji powtarzalnych procedur.

Ze względu na wczesne stadium w jakim znajduje się niniejsza koncepcja, niezbędne jest dokładniejsze porównanie jej z opublikowanymi już pracami o podobnej tematyce.

Następnym krokiem będzie stworzenie spójnego modelu pokrywającego wszystkie obszary zarządzania projektami zdefiniowane przez PMBOK [4], a następnie przekształcenie koncepcji w swoistą metametodologię umożliwiającą łączenie w jednym środowisku projektowych różnorodnych metodyk zarządzania projektami.

Istotnym punktem prac powinno być również opracowanie modelowego repozytorium obiektów, będącym punktem wyjścia przy implementacji opisywanego podejścia, oraz opracowanie koncepcji zestawu narzędzi informatycznych wspierających powyższe podejście.

Literatura

1. Fan S.L., Tserng H.P., Wang M.T. (2003). Development of an Object-Oriented Scheduling Model for Construction Projects. *Automation in Construction*, 12, s. 283-302.
2. Fan S.L., Tserng H.P. (2006). Object-Oriented Scheduling for Repetitive Projects with Soft Logics. *Journal of Construction Engineering and Management*, styczeń.
3. Karim A., Adeli H. (1999). OO Information Model for Construction Project Management. *Journal of Construction Engineering and Management*, 125(5), s. 361-367.
4. PMBOK Guide 2000 Edition, Kompendium wiedzy o zarządzaniu projektami. (2003). Management Training & Development Center.
5. Thayer R. (1999). *Visual Basic 6 Księga Eksperta*. Gliwice.
6. Tync Ł. (2005). *Podejście obiektowe w zarządzaniu projektami*. W: *Wybrane zagadnienia zarządzania projektami*. Red. T. Trzaskalik. AE, Katowice.

OBJECT ORIENTED APPROACH AND OBSERVER PATTERN SUPPORTING PROJECT COMMUNICATION AND RISK MANAGEMENT

Summary

It is quite obvious that modern global organizations are involved in many sophisticated projects and programs. Very often those projects are connected by various relations or constrains. Moreover it is hard to find a project, not affected by its environment nor other projects.

Increasing number of concurrent projects, high complexity and increasing group of stakeholders related to each specific project makes the project environment more and more complicated. In many situations, project is realized by few different teams, possible from different organizational culture. They use various methodologies and tools to assure highest level of control. This kind of heterogeneous environment is getting very common in multinational corporations.

In presented approach, it is assumed that there are certain analogies between project management and computer programming. On that basis, set of guidelines/rules can be transferred from object oriented programming and design patterns to project management. These rules can be used to order even very sophisticated structures.

This article shall give an overview of an object oriented approach and possibilities of its implementation in heterogeneous project environment. It is focused specially on design pattern called observer and the way how it can support information flow and risk control in the project or program.

RYZYKO W UBEZPIECZENIACH

Stanisław Heilpern

Uniwersytet Ekonomiczny we Wrocławiu

PROCESY RYZYKA Z ZALEŻNYMI WYPŁATAMI*

Wprowadzenie

W klasycznych modelach ryzyka [4, s. 156] autorzy przyjmowali wygodne, z matematycznego punktu widzenia, założenie mówiące o niezależności występujących w modelach zmiennych losowych i procesów. W naszej pracy odstępimy od tego założenia i dopuszczamy zależność wypłat. W praktyce wielkości wypłat są dość często zależne. Przykładowo, mogą istnieć czynniki zewnętrzne, takie jak ekstremalne zjawiska pogodowe, pożary, trzęsienia ziemi, karambole na autostradach czy kryzysy polityczne wpływające jednocześnie na wysokość wszystkich wypłat. Czynniki te powodują zależność wypłat.

W poprzednich pracach [2, s. 11; 3] przeprowadzona została analiza wpływu stopnia zależności wypłat na prawdopodobieństwo ruiny. Okazało się, że zależność ta nie jest regularna i w istotny sposób zależy od wartości kapitału początkowego. W naszej pracy zbadamy zależność prawdopodobieństwa ruiny od intensywności napływu składki.

1. Proces ryzyka

W pracy będziemy rozpatrywać ciągły proces ryzyka postaci [4, s. 156]

$$U(t) = u + ct - S(t)$$

gdzie u jest kapitałem początkowym, c intensywnością napływu składki, a $S(t) = \sum_{i=1}^{N(t)} X_i$ jest procesem opisującym zagregowane wypłaty do chwili t . W naszym modelu przyjmujemy, że wielkości wypłat $X_1, X_2, \dots$ są zmiennymi losowymi o jednakowym rozkładzie z wartością oczekiwaną m i dystrybuantą $F(x)$. Ponadto założymy, że proces liczby wypłat $\{N(t), t \geq 0\}$ jest procesem Poissona z intensywnością λ i jest niezależny od wielkości wypłat X_i .

* Praca naukowa finansowana ze środków na naukę w latach 2010-2012 jako projekt badawczy nr 3361/B/H03/2010/38.

Analizując powyższy model interesować nas będzie przede wszystkim prawdopodobieństwo zajścia ruiny, czyli zdarzenia, gdy wartość procesu $U(t)$ będzie po raz pierwszy ujemna. Rozpatrywać będziemy prawdopodobieństwo ruiny $\psi(u)$ w nieskończonym horyzoncie czasowym dla ustalonej wartości kapitału początkowego u

$$\psi(u) = P(T < \infty | U(0) = u)$$

gdzie $T = \inf\{t: U(t) < 0\}$ jest czasem wystąpienia ruiny.

W naszej pracy odejdziemy od klasycznego założenia teorii ryzyka, że wielkości wypłat są niezależne [4]. Dopuszczymy bardziej realistyczną sytuację, gdy wypłaty X_i mogą być zależne. Prawdopodobieństwo ruiny w przypadku niezależnych wypłat będziemy oznaczać symbolem $\psi_\lambda(u)$. Z klasycznej teorii ryzyka wiemy, że gdy $c \leq \lambda m$, to $\psi_\lambda(u) = 1$ dla każdego u , czyli zajście ruiny jest zdarzeniem pewnym. Ponadto dla skrajnych wartości kapitału początkowego otrzymujemy

$$\psi_\lambda(0) = \frac{\lambda}{c} m, \quad \psi_\lambda(\infty) = \lim_{u \rightarrow \infty} \psi_\lambda(u) = 0.$$

Drugim skrajnym przypadkiem jest całkowita zależność wielkości wypłat X_i . W tej sytuacji dla każdego i mamy tę samą zmienną losową X , a prawdopodobieństwo ruiny $\psi_c(u)$ określone jest wzorem [2, s. 14]

$$\psi_c(u) = \int_0^\infty \psi_x(u) dF(x) = \int_0^{c/\lambda} \psi_x(u) dF(x) + \bar{F}\left(\frac{c}{\lambda}\right)$$

gdzie $\psi_x(u)$ jest prawdopodobieństwem ruiny dla deterministycznych wypłat przyjmujących wartość $x > 0$, tzn. $P(X_i = x) = 1$. Gdy $x > c/\lambda$, to otrzymujemy $\psi_x(u) = 1$, a $\bar{F}(x) = 1 - F(x)$ jest funkcją przetrwania wypłat X_i . Prawdopodobieństwo ruiny $\psi_x(u)$ możemy obliczyć korzystając z kombinatorycznego wzoru [2]

$$\psi_x(u) = \begin{cases} 1 & \text{gdy } x \geq \frac{c}{\lambda} \\ 1 - \left(1 - \frac{\lambda}{c}\right) \sum_{k=0}^{\lfloor u/x \rfloor} \frac{1}{k!} (-z)^k e^z & \text{gdy } x < \frac{c}{\lambda} \end{cases}$$

gdzie $z = \frac{\lambda}{c}(u - kx)_+$, a $\lfloor w \rfloor$ jest największą liczbą naturalną nie większą niż w . Ponadto otrzymujemy zależność $\psi_x(0) = \lambda x / c$.

Gdy wypłaty są z dodatnim prawdopodobieństwem większe od c/λ , to wzajemna relacja między prawdopodobieństwami ruiny dla przypadku wypłat niezależnych i ściśle zależnych zależy istotnie od wielkości kapitału początkowego u . Fakt ten wynika z następujących zależności [2]

$$\psi_c(0) \leq \psi_I(0), \quad \psi_c(\infty) \geq \psi_I(\infty)$$

Gdy $F(\frac{c}{\lambda}) < 1$, to otrzymujemy ostre nierówności.

Dla małych wartości kapitału początkowego u mniejsze prawdopodobieństwo ruiny otrzymujemy dla ściśle zależnych wypłat. Natomiast dla bardzo dużych u „lepszym” przypadkiem, dającym mniejsze prawdopodobieństwo ruiny, są niezależne wypłaty.

Pośrednie stopnie zależności wypłat możemy modelować za pomocą funkcji łączących (copula). Funkcja łącząca C jest w naszej pracy łącznikiem między łączną funkcją przetrwania $\bar{G}(x_1, \dots, x_n) = P(X_1 > x_1, \dots, X_n > x_n)$, a brzegowymi funkcjami przetrwania $\bar{F}_i(x) = P(X_i > x_i)$ [6, s. 28]

$$\bar{G}(x_1, \dots, x_n) = C(\bar{F}_1(x_1), \dots, \bar{F}_n(x_n))$$

Archimedesowe funkcje łączące mają prostą, quasi-addywną postać indukowaną generatorem g [6, s. 89]

$$C(u_1, \dots, u_n) = g^{-1}(g(u_1) + \dots + g(u_n))$$

Generator $g: (0, 1] \rightarrow \mathbb{R}_+$ jest malejącą, ciągłą funkcją taką, że $\lim_{u \rightarrow 0} g(u) = \infty$ i $g(1) = 0$. Jeśli g^{-1} jest całkowicie monotoniczną funkcją, to jest transformatą Laplace’a pewnej nieujemnej zmiennej losowej Θ o dystrybucie F_Θ [6, s. 106; 1, s. 79]. Dla ustalonej wartości θ zmiennej losowej Θ zmienne X_i są warunkowo niezależne. Własność ta umożliwia stosowanie wtedy dla danej wartości θ znanych metod klasycznej teorii ruiny, a następnie wyznaczanie żądanej wartości obliczając odpowiednią całkę opartą na mieszanice indukowanej przez zmienną Θ . Zmienna ta generuje warunkowe zmienne losowe $X_{i|\theta}$ o funkcjach przetrwania [6, 1]

$$\bar{F}(x|\theta) = P(X_{i|\theta} > x) = \exp(-\theta g(\bar{F}(x)))$$

i wartościach oczekiwanych

$$m(\theta) = E(X_{i|\theta}) = \int_0^\infty \bar{F}(x|\theta) dx$$

oraz warunkowy proces ryzyka

$$U_\theta(t) = u + ct - S_\theta(t)$$

gdzie $S(t) = \sum_{i=1}^{N(t)} X_{i|\theta}$. Niech $\psi_\theta(u)$ oznaczać będzie warunkowe prawdopodobieństwo ruiny dla ustalonej wartości θ . Wtedy bezwarunkowe prawdopodobieństwo ruiny jest równe [2, s. 16]

$$\psi(u) = \int_0^{\infty} \psi_{\theta}(u) dF_{\Theta}(\theta) = \int_{\theta_0}^{\infty} \psi_{\theta}(u) dF_{\Theta}(\theta) + F_{\Theta}(\theta_0) \quad (1)$$

gdzie θ_0 jest największym θ spełniającym nierówność $m(\theta) \geq \frac{c}{\lambda}$. Funkcja $m(\theta)$ jest nierosnąca, więc dla każdego $\theta \leq \theta_0$ mamy $\psi_{\theta}(u) = 1$.

Tak jak w przypadku ścisłej zależności, można i w tym przypadku przeprowadzić porównanie z procesem ryzyka o niezależnych wypłatach. Zachodzą wtedy następujące zależności [2]

$$\psi(0) \leq \psi_I(0), \quad \psi(\infty) \geq \psi_I(\infty)$$

Otrzymujemy również te same relacje jak i w poprzednim przypadku, czyli brak regularności. Przypadek niezależnych wypłat jest „gorszy” niż zależne wypłaty, w których struktura zależności opisana jest archimedesową funkcją łączącą, dla małego kapitału początkowego u i „lepszy” dla bardzo dużego kapitału.

Indukowaną zmienną losową Θ możemy interpretować jako wpływ zewnętrznych czynników: makroekonomicznych, klimatycznych czy politycznych, oddziałujących w jednakowym stopniu na wszystkie wypłaty.

Archimedesowe funkcje łączące są dość proste i wygodne w stosowaniu. Jednak z samej konstrukcji tych funkcji wynika, że poszczególne pary zmiennych X_i, X_j są jednakowo skorelowane [6, s. 130]. Jest to niestety duże uproszczenie rzeczywistej sytuacji. W pracy archimedesowe funkcje łączące wykorzystane będą jedynie w celu pokazania, w jaki sposób intensywność napływu składki oraz inne czynniki wpływają na prawdopodobieństwo ruiny i jakie nieregularności możemy wtedy zaobserwować.

Przyjmijmy też dalsze uproszczenia, umożliwiające wygodniejsze wyznaczenie prawdopodobieństwa ruiny. Załóżmy, że zmienne losowe X_i są dyskretne i przyjmują jedynie dwie wartości: x_1 z prawdopodobieństwem p oraz x_2 z prawdopodobieństwem $q = 1 - p$ [3]. Odpowiada to sytuacji, gdy wypłaty są podzielone na dwie istotnie różne, jednorodne kategorie. Wtedy prawdopodobieństwo ruiny można wyznaczyć dla niezależnych wypłat wykorzystując kombinatoryczny wzór [4, s. 102]

$$\psi_I(u) = \begin{cases} 1 & \text{gdy } x \geq \frac{c}{\lambda} \\ 1 - \left(1 - \frac{\lambda}{c}\right) \sum_{k_1=0}^{\lfloor \frac{u}{x_1} \rfloor} \sum_{k_2=0}^{\lfloor \frac{u-k_1x_1}{x_2} \rfloor} \frac{p^{k_1} q^{k_2}}{k_1! k_2!} (-z)^{k_1+k_2} e^z & \text{gdy } x < \frac{c}{\lambda} \end{cases} \quad (2)$$

gdzie $z = \frac{\lambda}{c}(u - k_1x_1 - k_2x_2)_+$.

W przypadku ściśle zależnych wypłat prawdopodobieństwo ruiny określone jest prostym wzorem

$$\psi_c(u) = p\psi_{x_1}(u) + q\psi_{x_2}(u) \quad (3)$$

Gdy struktura zależności wypłat opisana jest archimedesową funkcją łączącą, to warunkowe zmienne losowe $X_{i|\theta}$ również mają dla ustalonej wartości indukowanej zmiennej losowej Θ rozkład dwupunktowy i przyjmują wartość x_1 z prawdopodobieństwem $p(\theta) = 1 - \exp(-\theta g(q))$ oraz x_2 z prawdopodobieństwem $q(\theta) = 1 - p(\theta)$. Natomiast graniczna wartość zmiennej Θ jest równa [3]

$$\theta_0 = -\frac{1}{g(q)} \ln \left(\frac{c - \lambda x_1}{\lambda(x_1 - x_2)} \right)$$

W przedstawionych w dalszej części pracy przykładach wykorzystywać będziemy archimedesową funkcję łączącą Claytona

$$C_\alpha(u_1, \dots, u_n) = (u_1^{-\alpha} + \dots + u_n^{-\alpha})^{-1/\alpha}$$

gdzie $\alpha > 0$. Parametr α wyznacza w tym przypadku stopień zależności zmiennych losowych X_i . Współczynnik korelacji τ Kendalla jest wtedy równy [6, s. 130]

$$\tau = \frac{\alpha}{\alpha + 2}$$

W granicznym przypadku $\alpha = 0$ otrzymujemy niezależność, a dla $\alpha = \infty$ ścisłą zależność. Funkcja łącząca Claytona indukuje zmienną losową Θ o rozkładzie gamma $\text{Ga}(1/\alpha, 1)$ z dystrybuantą F_α . Jej generatorem jest funkcja $g(u) = u^{-\alpha} - 1$, a warunkowe prawdopodobieństwo otrzymania wartości wypłaty równej x_2 wynosi $q(\theta) = \exp(\theta(1 - q^{-\alpha}))$.

W dalszej części pracy przyjmiemy dodatkowe, upraszczające założenia. Założymy, że intensywność występowania wypłat $\lambda = 1$, tzn. należy się spodziewać przeciętnie jednej wypłaty w jednostce czasu, $x_1 = 1$ oraz $x_2 = 2$. Wtedy $m(\theta) = 1 + q(\theta)$, a graniczna wartość indukowanej zmiennej Θ zależy od wartości parametru α i przyjmuje prostą postać

$$\theta_\alpha = \frac{\ln(c-1)}{1-q^{-\alpha}}$$

Dla ustalonej wartości parametru α możemy obliczyć prawdopodobieństwo ruiny $\psi_\alpha(u)$ dla skrajnych wartości kapitału początkowego korzystając ze wartości dystrybuanty F_α rozkładu gamma

$$\psi_\alpha(0) = \int_{\theta_\alpha}^{\infty} \frac{m(\theta)}{c} dF_\alpha(\theta) + F_\alpha(\theta_\alpha) = \frac{1}{c}(\bar{F}_\alpha(\theta_\alpha) + q\bar{F}_\alpha(q^{-\alpha}\theta_\alpha)) + F_\alpha(\theta_\alpha)$$

oraz $\psi_\alpha(\infty) = F_\alpha(\theta_\alpha)$. Natomiast wzory (1) i (2) umożliwiają nam obliczenie tych prawdopodobieństw dla dowolnych wartości kapitału początkowego u .

2. Analiza zależności prawdopodobieństwa ruiny od intensywności napływu składki

Prawdopodobieństwo ruiny zależy od stopnia zależności wypłat. Jak już zostało wcześniej wspomniane, zależność ta jest jednak dość nieregularna, zależy istotnie od wartości kapitału początkowego u [2, s. 11; 3].

Zbadamy teraz w jakim stopniu wielkość składki c wpływa na prawdopodobieństwo ruiny. Wiemy (patrz 2), że jeśli dla niezależnych wypłat zachodzi $c \leq m$ (obowiązuje założenie $\lambda = 1$), to $\psi_c(u) = 1$. W przypadku ściśle zależnych wypłat, korzystając z (3) otrzymujemy

$$\psi_c(u) = \begin{cases} 1 & \text{gdy } c \leq 1 \\ p\psi_1(u) + q & \text{gdy } 1 < c \leq 2 \\ p\psi_1(u) + q\psi_2(u) & \text{gdy } 2 < c \end{cases}$$

Natomiast dla wypłat o strukturze zależności opisaną funkcją łączącą Claytona graniczna wartość zmiennej Θ określona jest wzorem

$$\theta_{\alpha} = \begin{cases} \infty & \text{gdy } c \leq 1 \\ \frac{\ln(c-1)}{1-q^{-\alpha}} & \text{gdy } 1 < c \leq 2 \\ 0 & \text{gdy } 2 < c \end{cases}$$

Wynika stąd oraz z (1), że dla $c \leq 1$ zachodzi $\psi_{\alpha}(u) = 1$ dla każdego u i α .

Przeanalizujemy teraz przypadki wyznaczone przez skrajne wartości kapitału początkowego u rozpatrując wartości parametru α oddającego stopień zależności wypłat równe 0; 0,1; 1; 2; 4 oraz ∞ . Odpowiadają im odpowiednio wartości współczynnika korelacji τ Kendalla równe 0; 0,048; 0,333; 0,5; 0,667 oraz 1. Zaczniemy od zerowego kapitału początkowego.

Przykład 1

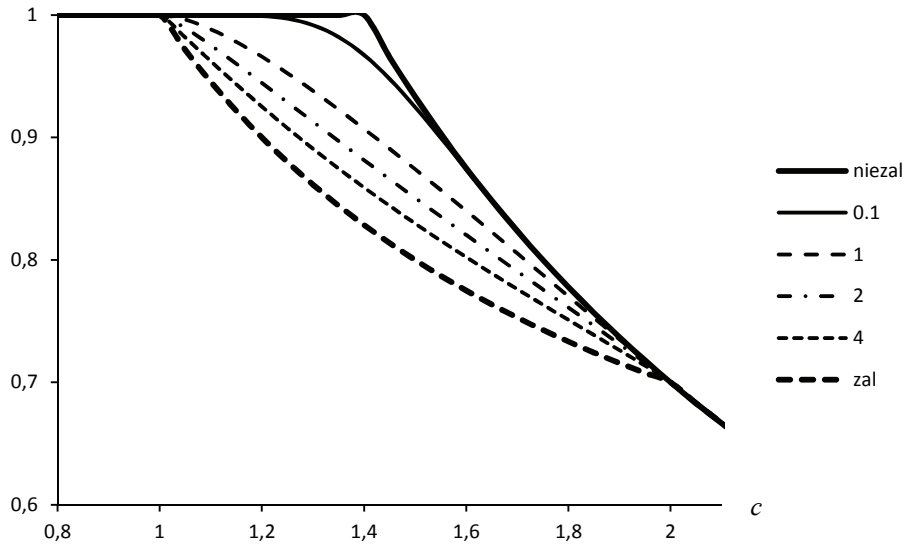
Niech $u = 0$ oraz $q = 0,4$. W tabeli 1 i na rys. 1 przedstawione są wartości prawdopodobieństw ruiny.

Tabela 1

Prawdopodobieństwo ruiny dla $u = 0$ i różnych c oraz α

c	α					
	niezależne	0,1	1	2	4	ściśle zal.
0	1	1	1	1	1	1
1,0	1	1	1	1	1	1
1,1	1	0,999995	0,988249	0,97526	0,962211	0,945455
1,2	1	0,999267	0,965800	0,944492	0,925101	0,900000
1,3	1	0,991793	0,937950	0,912845	0,890765	0,861538
1,4	1	0,967361	0,906934	0,881435	0,859085	0,828571
1,5	0,933333	0,924800	0,874008	0,850594	0,829688	0,800000
1,6	0,875000	0,873676	0,839940	0,820356	0,802163	0,775000
1,7	0,823529	0,823435	0,805225	0,790599	0,776088	0,752941
1,8	0,777778	0,777776	0,770194	0,76107	0,750998	0,733333
1,9	0,736842	0,736842	0,735068	0,731314	0,726253	0,715789
2,0	0,700000	0,700000	0,700000	0,700000	0,700000	0,700000
2,5	0,560000	0,560000	0,560000	0,560000	0,560000	0,560000
3,0	0,466667	0,466667	0,466667	0,466667	0,466667	0,466667
∞	0	0	0	0	0	0

Widzimy, że dla zerowego kapitału początkowego oraz dla wartości intensywności napływu składki c zawartej w przedziale $(1; 2)$ prawdopodobieństwo ruiny maleje wraz ze stopniem zależności wypłat. Natomiast dla $c \geq 2$ oraz $c \leq 1$ prawdopodobieństwo to nie zależy od wartości α . W drugim przypadku jest zawsze równe 1.



Rys. 1. Prawdopodobieństwo ruiny dla $u = 0$ i różnych c oraz α .

Inną, bardziej skomplikowaną, nieregularną sytuację otrzymujemy dla nieskończenie dużego kapitału początkowego.

Przykład 2

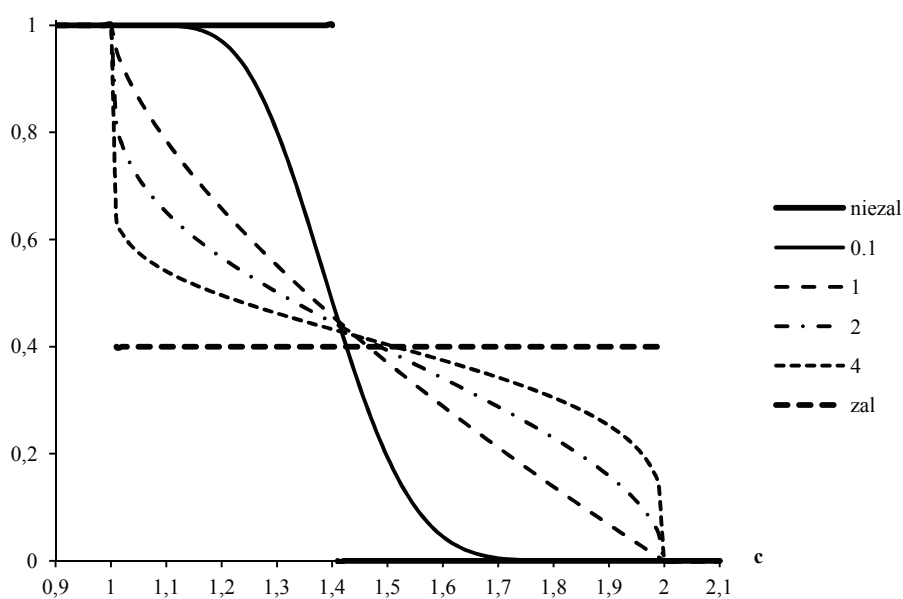
Niech $u = \infty$ oraz $q = 0,4$. W tabeli 2 zawarte są wartości prawdopodobieństw ruiny oraz przedstawione na rys. 2.

Dla intensywności napływu składki c mniejszej od wartości oczekiwanej wypłat 1,4 prawdopodobieństwo ruiny zależy w sposób regularny, monotoniczny od stopnia zależności wypłat. Prawdopodobieństwo maleje wraz ze wzrostem parametru α . Dla wartości $c > 1,52$ obserwujemy relację odwrotną. Prawdopodobieństwo ruiny rośnie, gdy rośnie stopień zależności wypłat.

Tabela 2

Prawdopodobieństwo ruiny dla $u = \infty$ i różnych c oraz α

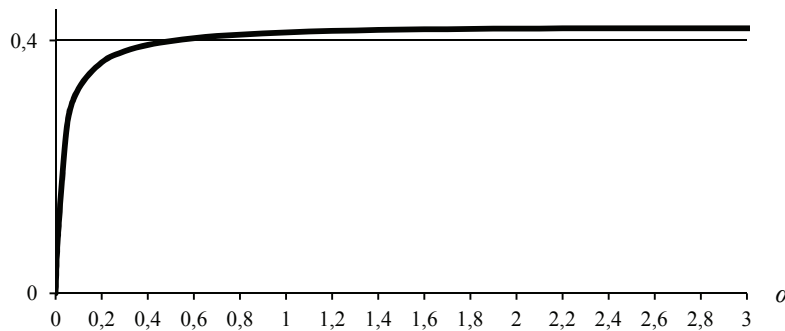
c	α					
	niezależne	0.1	1	2	4	ściśle zal.
0	1	1	1	1	1	1
1.00	1	1	1	1	1	1
1.10	1	0,999573	0,784557	0,651023	0,540642	0,4
1.20	1	0,970624	0,658005	0,566384	0,496110	0,4
1.30	1	0,802130	0,551860	0,501747	0,462356	0,4
1.40	1	0,484516	0,457116	0,445356	0,432495	0,4
1.45	0	0,323970	0,412770	0,418735	0,418128	0,4
1.50	0	0,192864	0,370039	0,392653	0,403817	0,4
1.60	0	0,045128	0,288621	0,340885	0,374507	0,4
1.70	0	0,005001	0,211626	0,287584	0,342618	0,4
1.80	0	0,000157	0,138226	0,229377	0,304924	0,4
1.90	0	0,000000	0,067830	0,158788	0,252920	0,4
2.00	0	0	0	0	0	0
∞	0	0	0	0	0	0

Rys. 2. Prawdopodobieństwo ruiny dla $u = \infty$ i różnych c oraz α

Natomiast dla intensywności c zawartej między 1,4 a 1,52 nie ma już regularnej, monotonicznej zależności. W tabeli 3 oraz na rys. 3 przedstawione są wartości prawdopodobieństwa ruiny gdy $c = 1,45$ dla różnych wartości parametru α .

Tabela 3

Prawdopodobieństwo ruiny dla $u = \infty$, $c = 1,45$ i różnych α										
α	0	0,1	0,2	1	2	3	5	10	20	∞
$\psi_\alpha(\infty)$	0	0,3240	0,3654	0,4128	0,4187	0,4191	0,4168	0,4111	0,4063	0,4

Rys. 3. Prawdopodobieństwo ruiny dla $u = \infty$, $c = 1,45$ i różnych α

Jak można zauważyć, w tym przypadku nie ma monotoniczności. Dla małych wartości parametru α prawdopodobieństwo ruiny rośnie. Maksimum równe 0,41924 osiąga około $\alpha = 2,628$ ($\tau = 0,568$), a następnie dla większych wartości α lekko spada, dążąc w nieskończoność do wartości osiągalnej dla ściśle zależnych wypłat $\psi_c(\infty) = 0,4$. Należy też zaznaczyć, że stabilizacja wartości prawdopodobieństwa ruiny wokół 0,4 zachodzi już dużo wcześniej, dla wartości α większych od 0,4 ($\tau = 0,167$). Dla mniejszych wartości mamy gwałtowny wzrost prawdopodobieństwa ruiny.

Podobny efekt możemy również zaobserwować dla pośrednich wartości kapitału początkowego u .

Przykład 3

Niech $u = 24$ oraz $q = 0,4$. Wartości prawdopodobieństwa ruiny zamieszczone są w tabeli 4 oraz przedstawione na rys. 4.

W tym przypadku również występuje obszar braku monotoniczności. Sytuacja ta zachodzi gdy wartości intensywności napływu składki c należą do przedziału $[1,4; 1,6)$. W odróżnieniu od przypadku dla nieskończonego kapitału początkowego, wykresy funkcji prawdopodobieństwa ruiny są ciągłe. Ponadto występują dodatnie wartości prawdopodobieństwa ruiny dla $c > 2$.

Tabela 4

Prawdopodobieństwo ruiny dla $u = 24$ i różnych c oraz α

c	α					
	niezależne	0,1	1	2	4	ściśle zal.
1,0	1	1	1	1	1	1
1,1	1	0,999855	0,823578	0,684144	0,561616	0,406276
1,2	1	0,984932	0,694453	0,590099	0,508586	0,400104
1,3	1	0,874290	0,589201	0,524356	0,474183	0,400004
1,4	1	0,618220	0,496195	0,468649	0,444862	0,400000
1,5	0,071154	0,324054	0,411150	0,417618	0,417433	0,400000
1,6	0,006457	0,122893	0,331878	0,368447	0,390128	0,400000
1,7	0,000719	0,034607	0,257069	0,318984	0,361383	0,400000
1,8	0,000096	0,008200	0,185860	0,266798	0,329054	0,400000
1,9	0,000015	0,001974	0,117643	0,207608	0,288597	0,400000
2,0	0,000003	0,000522	0,051971	0,123987	0,215211	0,400000
2,2	0,000000	0,000047	0,005059	0,012157	0,021181	0,039515
2,5	0,000000	0,000002	0,000248	0,000601	0,001051	0,001970
3,0	0,000000	0,000000	0,000004	0,000010	0,000018	0,000033
∞	0	0	0	0	0	0

W tabeli 5 oraz na rys. 5 przedstawione są wartości prawdopodobieństwa ruiny dla kapitału początkowego $u = 24$ i różnych stopni zależności wypłat.

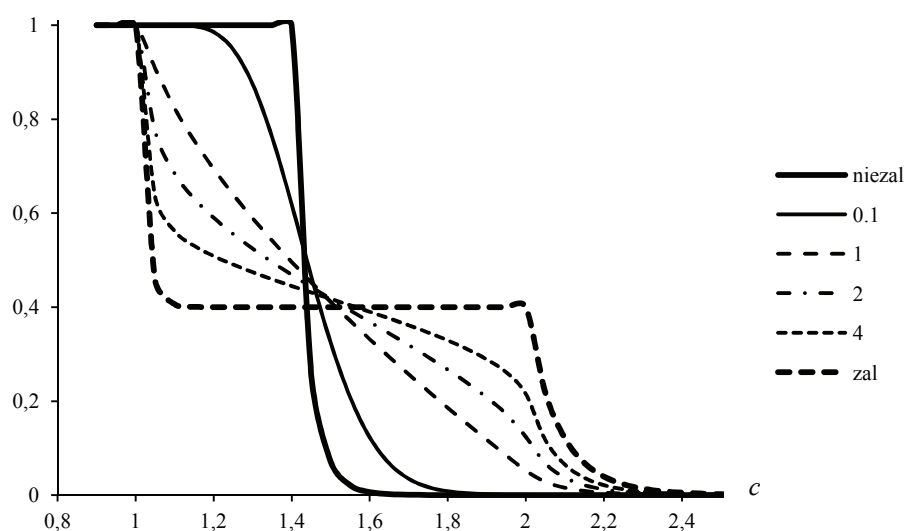
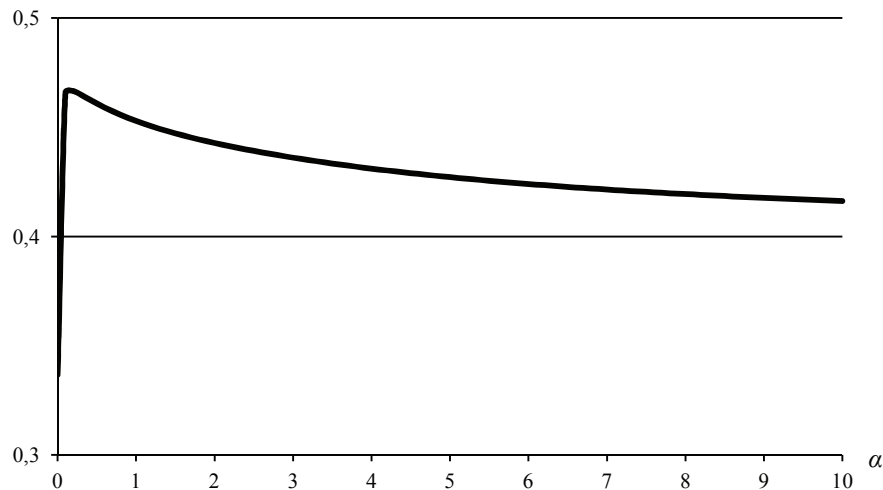
Rys. 4. Prawdopodobieństwo ruiny dla $u = 24$ i różnych c oraz α

Tabela 5

Prawdopodobieństwo ruiny dla $u = 24$, $c = 1,45$ i różnych α

α	0	0,1	0,2	0,5	1	2	3	5	10	∞
$\psi_\alpha(24)$	0,3366	0,4658	0,4665	0,4608	0,4528	0,4428	0,4361	0,4271	0,4162	0,4

Rys. 5. Prawdopodobieństwo ruiny dla $u = 24$, $c = 1,45$ i różnych α

Widzimy, że dla bardzo małych wartości parametru α prawdopodobieństwo ruiny gwałtownie rośnie i osiąga maksimum równe 0,4668 dla $\alpha_0 = 0,16$ ($\tau = 0,074$). Następnie spokojnie maleje do wartości 0,4 osiągalnej dla ściśle zależnych wypłat. Dla małych wartości stopnia zależności wypłat, bliskich niezależności, funkcja prawdopodobieństwa ruiny $\psi_\alpha(24)$, traktowana jako funkcja parametru α , jest bardzo wrażliwa na zmianę wartości argumentu. Na szczęście, w praktyce, w większości przypadków stopień zależności raczej przekracza α_0 i otrzymujemy wtedy monotoniczną, słabo malejącą tendencję.

Na zakończenie zbadamy przypadek, gdy intensywność napływu składki c jest znaczna i przekracza x_2 .

Tabela 6

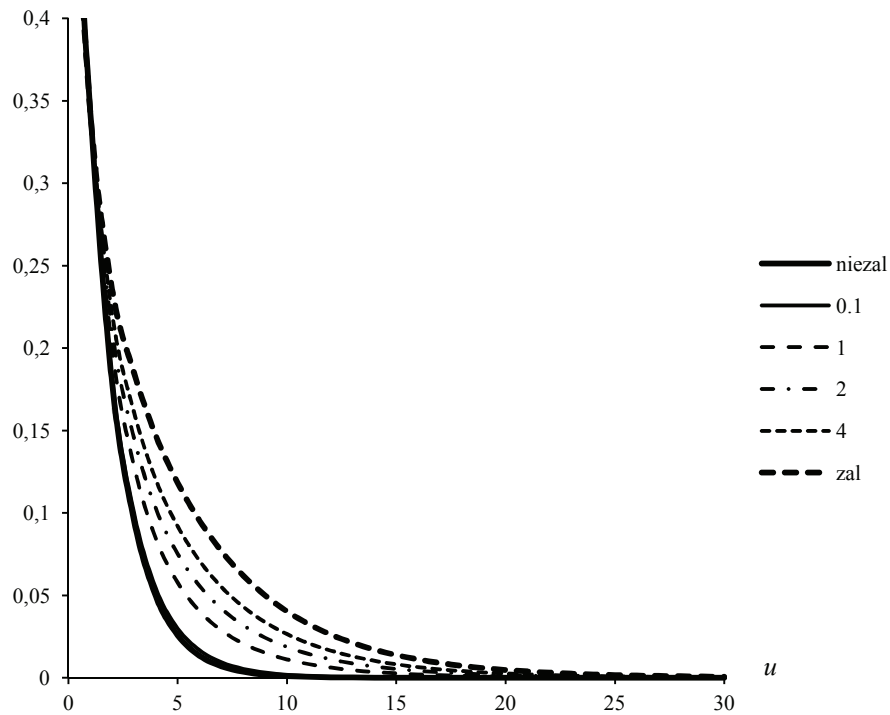
Prawdopodobieństwo ruiny dla $c = 2.1$, różnych u oraz α

c	α					
	niezależne	0,1	1	2	4	ściśle zal.
0	0,560000	0,560000	0,560000	0,560000	0,560000	0,560000
1	0,343597	0,343597	0,343597	0,343597	0,343597	0,343597
2	0,178299	0,181354	0,199781	0,210494	0,220653	0,235585
3	0,095303	0,099972	0,128503	0,145240	0,161188	0,184878
4	0,049333	0,054123	0,084397	0,102564	0,120079	0,146737
5	0,025703	0,029891	0,057752	0,075010	0,091879	0,118167
6	0,013342	0,016654	0,040369	0,055689	0,070909	0,095181
7	0,006931	0,009397	0,028744	0,041859	0,055112	0,076716
8	0,003600	0,005361	0,020753	0,031745	0,043048	0,061851
9	0,001870	0,003091	0,015150	0,024237	0,033743	0,049861
10	0,000972	0,001800	0,011160	0,018605	0,026527	0,040199
15	0,000037	0,000138	0,002656	0,005233	0,008193	0,013691
20	0,000001	0,000013	0,000695	0,001553	0,002602	0,004663
25	0,000000	0,000001	0,000192	0,000475	0,000839	0,001588
30	0,000000	0,000000	0,000055	0,000148	0,000273	0,000541
∞	0	0	0	0	0	0

Przykład 4

Niech $c = 2,1$ i $q = 0,4$. Wyznamy prawdopodobieństwo ruiny dla różnych wartości kapitału początkowego u i wybranych wartości parametru α równych 0; 0,1; 1; 2; 4 i ∞ . Wyniki przedstawione są w tabeli 6 i na rys. 6.

W odróżnieniu od poprzednich przykładów, w tym przypadku otrzymaliśmy regularny, monotoniczny przypadek. Prawdopodobieństwo ruiny wzrasta wraz ze stopniem zależności wypłat dla każdej wartości kapitału początkowego. Wartości prawdopodobieństwa ruiny nie są też tak wrażliwe na zmiany wartości parametru α jak w poprzednim przykładzie. Przykładowo, dla $\alpha = 0,1$ prawdopodobieństwa te są niewiele większe niż dla niezależności (rys. 6).



Rys. 6. Prawdopodobieństwo ruiny dla $c = 2, 1$; różnych u oraz α

Podsumowanie

W pracy przedstawiona została analiza wpływu intensywności napływu składki c na prawdopodobieństwo ruiny. Rozpatrzono warianty obejmujące różne wartości kapitału początkowego u i stopnia zależności wypłat reprezentowanego poprzez wartość parametru α rodziny funkcji łączących Clayтона. Otrzymano w niektórych przypadkach nieregularność, brak monotonicznej zależności prawdopodobieństwa ruiny od tych czynników. W następnych pracach autor zamierza zbadać wpływ prawdopodobieństwa q wystąpienia większej wypłaty na prawdopodobieństwo ruiny.

Literatura

1. Genest C., Marceau E., Mesfioui M. (2003). Compound Poisson Approximations for Individual Models with Dependent Risks. *Insurance: Mathematics and Economics*, 32.
2. Heilpern S. (2010). Wyznaczanie prawdopodobieństwa ruiny, gdy struktura zależności wypłat jest opisana archimedesową funkcją łączącą. W: *Zagadnienia aktuarialne – teoria i praktyka*. Red. W. Otto. UW, Warszawa.
3. Heilpern S. Analiza wpływu stopnia zależności wypłat na prawdopodobieństwo ruiny. *Ekonometria* (w druku).
4. Kaas R., Goovaerts M., Dhaene J., Denuit M. (2001). *Modern Actuarial Risk Theory*. Kluwer, Boston.
5. *Modele aktuarialne*. (2000). Red. W. Ostasiewicz. AE, Wrocław.
6. Nelsen R.B. (1999). *An Introduction to Copulas*. Springer, New York.

RISK PROCESSES WITH DEPENDENT CLAIMS

Summary

The risk processes when the claims may be dependent are studied. It is more realistic attempt than in the classical risk process, when the independent claims are assumed. The continuous risk process is investigated from the ruin theory point of view. The probability of ruin within an infinite time horizon is derived. The influence of the premium rate on the probability of ruin is analysed. The dependent structure of claims is described by the Archimedean copula. The case of the strict dependent claims are studied, too. The nonregular, nonmonotonic relations between the probability of ruin and the premium rate or other factors are observed.

Rafał Korzonek

Uniwersytet Ekonomiczny we Wrocławiu

WPŁYW RYZYKA DŁUGOWIECZNOŚCI NA UBEZPIECZENIA NA ŻYCIE

Wprowadzenie

Przemiany demograficzne zachodzące w ciągu ostatnich dekad wskazują na postępujące starzenie się społeczeństwa, a tym samym powodują wydłużanie się średniej długości trwania życia. Jak wiadomo, jednym z podstawowych rodzajów ryzyka ubezpieczeniowego występującego w ubezpieczeniach na życie jest długość dalszego trwania życia osoby ubezpieczonej. Zakłady ubezpieczeń do kalkulacji składki netto wykorzystują dane zawarte w tablicach trwania życia. Tablice trwania życia zawierają informacje o prawdopodobieństwie zgonu w ciągu roku osoby w określonym wieku, liczbie osób dożywających określonego wieku, liczbie osób zmarłych w określonym wieku, przeciętnym dalszym trwaniu życia osób w określonym wieku. W wyniku zmian demograficznych ulegają one ciągłym zmianom. Zatem w konsekwencji możemy stwierdzić, że proces starzenia się populacji wpływa na wielkość i ryzyko związane z wyceną produktów ubezpieczeń na życie. Celem opracowania jest prezentacja zmian demograficznych oraz roli ryzyka długowieczności w procesie wyceny składki netto w ubezpieczeniach na życie.

1. Ryzyko długowieczności

W literaturze dotyczącej aspektów demograficznych można zauważyć znaczne rozbieżności dotyczące oczekiwanej długości ludzkiego życia. Na początek warto wspomnieć, że Olshansky, Carnes i Cassel twierdzą, iż nie można oczekiwać, aby długość ludzkiego życia wynosiła więcej niż 85 lat. Odmienne zdania są Oeppen i Vaupel, którzy przeprowadzili badania dotyczące długości życia kobiet w XX wieku i stwierdzili, że wzrost długości życia kobiet jest stały i wynosi 2,5 roku na dekadę. Twierdzą oni także, że w niektórych wy-

soko rozwiniętych krajach już w 2060 roku długość ludzkiego życia może wynieść 100 lat. W tym samym tonie wypowiadają się Sanderson i Scherbov, stwierdzając, że średnia długość życia już zbliża się do 85 lat i nie ma żadnych podstaw statystycznych by sądzić, że istnieje bariera długowieczności.

W rzeczywistości analizując dane dotyczące umieralności można stwierdzić, że z roku na rok następuje poprawa wskaźników umieralności. Dzięki temu wydłuża się średnia długość dalszego trwania życia. Wśród podstawowych czynników, które wpływają na tę poprawę można wyróżnić wzrost gospodarczy, poprawę warunków życia, poprawę poziomu opieki medycznej, postęp medycyny, poprawę sytuacji materialnej, upowszechnianie aktywnego wypoczynku, upowszechnianie zdrowego odżywiania się.

Według prognoz przedstawionych przez Europejski Urząd Statystyczny, ludność Unii Europejskiej starzeje się i nadal będzie się starzeć. Eurostat podaje, że zwiększy się udział w populacji ludzi w wieku 65 lat i starszych z poziomu 17.1% w 2008 roku do 30.0% w 2060 oraz tych w wieku ponad 80 lat z 4.4% do 12.1% w tym samym okresie. Ludność UE będzie się starzeć w ciągu całego badanego okresu, szczególnie poprzez ustawicznie niski wskaźnik dzietności oraz rosnącą liczbę osób dożywających starości. Proces starzenia wystąpi we wszystkich państwach członkowskich. W 2060 roku udział ludności w wieku 65 lat lub większym wzrośnie do 36.2% w Polsce, 36.1% w Słowacji i 35.0% w Rumunii. W konsekwencji stosunek ludności w wieku 65 lat i wyższym do ludności czynnej zawodowo, zwiększy się według przewidywań z 25% w 2008 roku do 53% w 2060. Innymi słowy, na każdą osobę w wieku 65 roku życia i powyżej będą przypadały w 2060 roku jedynie dwie osoby czynne zawodowo, w porównaniu do 4 czynnych zawodowo, przypadających na każdą osobę w starszym wieku obecnie.

Warto także przyjrzeć się aktualnej oczekiwanej średniej długości trwania. Według badań The World Factbook na 2007 rok spośród krajów członkowskich Unii Europejskiej najdłużej żyją obywatele Szwecji – średnio 80,63 lat. Tuż za nimi znajdują się obywatele Francji, Włoch, Hiszpanii i Grecji. Polska znajduje się w tym zestawieniu dopiero na 20 miejscu ze średnią długością życia wynoszącą 75,19 lat, przy czym średnia długość życia mężczyzn to 71,18 lat, a kobiet 79,44. Warto także zauważyć, że kraje sąsiedzkie mają bardzo zbliżoną do nas średnią długość życia. Według badań The World Factbook w 2007 roku najkrócej żyli mieszkańcy Estonii, Rumunii i Łotwy.

Tabela 1

Oczekiwana średnia długość trwania życia w krajach członkowskich UE

Lp.	Kraj	Ogół	Mężczyźni	Kobiety
1	Szwecja	80,63	78,39	83,00
2	Francja	80,59	77,35	84,00
3	Włochy	79,94	77,01	83,07
4	Hiszpania	79,78	76,46	83,32
5	Grecja	79,38	76,85	82,06
6	Austria	79,21	76,32	82,26
7	Malta	79,15	76,95	81,47
8	Holandia	79,11	76,52	81,82
9	Luksemburg	79,03	75,76	82,52
10	Niemcy	78,95	75,96	82,11
11	Belgia	78,92	75,75	82,24
12	Wielka Brytania	78,70	76,23	81,30
13	Finlandia	78,66	75,15	82,31
14	Cypr	77,98	75,60	80,49
15	Dania	77,96	75,65	80,41
16	Irlandia	77,90	75,27	80,70
17	Portugalia	77,87	74,60	81,36
18	Słowenia	76,53	72,84	80,47
19	Czechy	76,42	73,14	79,88
20	Polska	75,19	71,18	79,44
21	Słowacja	74,95	71,00	79,11
22	Litwa	74,44	69,46	79,69
23	Węgry	72,92	68,73	77,38
24	Bułgaria	72,57	68,95	76,40
25	Estonia	72,30	66,87	78,07
26	Rumunia	71,91	68,41	75,62
27	Łotwa	71,60	66,39	77,10

Źródło: The World Factbook 2007.

Tabela 2

Oczekiwana średnia długość życia w Polsce w latach 1950-2009

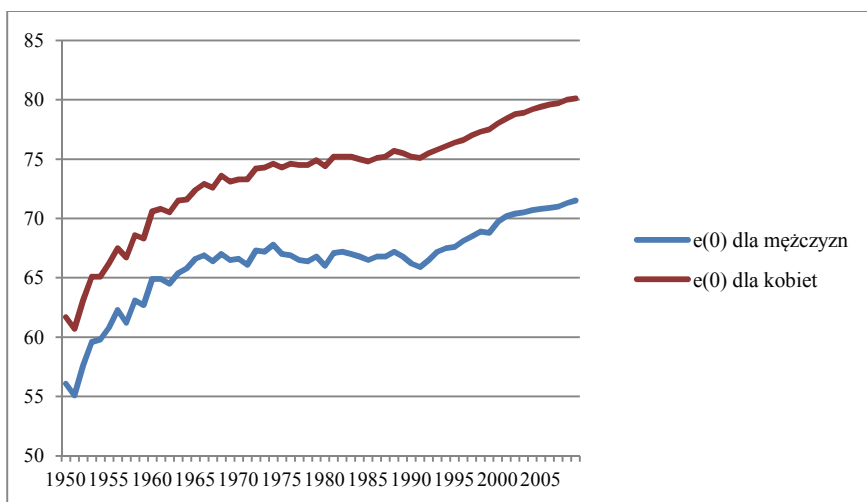
Lata	e(0) dla mężczyzn	e(0) dla kobiet
1	2	3
1950	56,1	61,7
1955	60,8	66,2
1960	64,9	70,6
1965	66,6	72,4
1970	66,6	73,3

cd. tabeli 2

1	2	3
1975	67,0	74,3
1980	66,0	74,4
1985	66,5	74,8
1990	66,2	75,2
1995	67,6	76,4
2000	69,7	78,0
2005	70,8	79,4
2009	71,5	80,1

Źródło: Dane GUS.

Bardzo ciekawie prezentuje się także zestawienie oczekiwanej średniej długości życia w Polsce w latach 1950-2009. Widać tutaj znaczną poprawę umieralności. W 2009 roku w Polsce mężczyźni żyli przeciętnie 71,5 lat, natomiast kobiety 80,1. W stosunku do 1950 roku mężczyźni żyją dłużej o 15,4 lat, natomiast kobiety o 18,4. Patrząc na wykres średniego dalszego trwania życia w Polsce można stwierdzić, że poziom umieralności w Polsce był bardzo wysoki bezpośrednio po II wojnie światowej. W 1950 roku przeciętna długość trwania życia mężczyzn wynosiła nieco ponad 56 lat, natomiast kobiet prawie 62 lata. Dzięki istotnemu postępowi w zmniejszeniu poziomu umieralności zarówno mężczyzn, jak i kobiet, a zwłaszcza dzięki wzmocnieniu istniejącego trendu spadku umieralności niemowląt, osiągnięto znaczny wzrost długości trwania życia.



Rys. 1. Średnie dalsze trwanie życia w Polsce w latach 1950-2009

Źródło: Opracowanie na podstawie danych GUS.

3. Ryzyko długowieczności a ubezpieczenia na życie

Jak wspomniano na wstępie jednym z podstawowych rodzajów ryzyka ubezpieczeniowego występującego w ubezpieczeniach na życie jest długość dalszego trwania życia osoby ubezpieczonej. Proces starzenia się społeczeństwa i związanego z tym ryzyka może powodować różne skutki finansowe dla zakładu ubezpieczeń. W zależności od rodzaju ubezpieczenia skutki te mogą mieć zarówno charakter pozytywny, jak i negatywny. Dla przykładu, w sytuacji ubezpieczeń renty dożywotniej coraz dłuższy czas życia osoby ubezpieczonej ma niekorzystny wpływ na finanse ubezpieczyciela. Suma zebranych składek plus zyski z inwestycji mogą okazać się niewystarczające na pokrycie zwiększonej liczby wypłat renty. Natomiast w przypadku dożywotniego ubezpieczenia na wypadek śmierci dłuższy czas życia osoby ubezpieczonej powoduje odroczenie w czasie wypłaty świadczenia, dzięki czemu ubezpieczyciel zbiera dodatkowe składki oraz osiąga dodatkowy zysk z inwestycji. Powstaje zatem pytanie, jak kalkulować składkę netto, aby uniknąć niedoszacowania przyszłych świadczeń jakie ponosi zakład ubezpieczeń? Jednym z rozwiązań wydaje się uwzględnienie zmieniającej się umieralności. Jak wiadomo, funkcjonujące na Polskim rynku zakłady ubezpieczeń na ogół korzystają z tablic trwania życia publikowanych przez Główny Urząd Statystyczny. Podstawą budowania pełnych tablic trwania życia są dane dotyczące przeszłości – liczby osób zmarłych w minionym roku, ponadto zakłada się, że roczna liczba urodzeń i zgonów, a także umieralność jest jednakowa i niezmienna w poszczególnych latach. Zatem przewidywanie przyszłości opiera się wyłącznie na warunkach jakie panowały w przeszłości. Aby zmniejszyć ryzyko niewłaściwego oszacowania wartości produktów i rezerw, należy wprowadzić do wyceny produktów ubezpieczeniowych zamiast tradycyjnych tablic trwania życia tzw. dynamiczne tablice trwania życia, które uwzględnią informacje o zmieniającej się umieralności. Jednym z modeli pozwalającym na budowanie dynamicznych tablic trwania życia jest model Lee i Cartera. Model ten uwzględnia efekt starzenia się populacji i być może dzięki temu jest jednym z najczęściej wykorzystywanych modeli w badaniach demograficznych. Niżej w artykule przedstawiono opis modelu Lee i Cartera zaczerpnięty z [6].

W latach 90. Lee i Carter podjęli próbę zastosowania teorii procesu błędzenia przypadkowego z dryfem, do modelowania oraz prognozowania współczynników zgonów $m_x(t)$ (*central death rate*) w grupach wieku indeksowanych zmienną x i dla kolejnych lat kalendarzowych indeksowanych zmienną t .

Model Lee i Cartera (1992) ma postać

$$\ln m_x(t) = a_x + b_x k_t + \varepsilon_{x,t} \quad \text{dla } x=0,1,\dots,\omega, t=1,2,\dots, \quad (1)$$

lub równoważnie

$$m_x(t) = \exp \{a_x + b_x k_t + \varepsilon_{x,t}\} \quad \text{dla } x=0,1,\dots,\omega, t=1,2,\dots, \quad (2)$$

gdzie:

$\{a_x\}$ oraz $\{b_x\}$ – są zbiorami pewnych stałych różnych dla różnych grup wieku x ,

k_t – jest indeksem, traktowanym jako pewien proces stochastyczny z czasem dyskretnym (opisuje ogólną tendencję zmian umieralności),

$\varepsilon_{x,t}$ – są składnikami losowymi, niezależnymi o rozkładach $N(0, \sigma^2)$.

Dla zapewnienia jednoznaczności rozwiązania konieczne jest określenie pewnych dodatkowych warunków ograniczających. W tym celu Lee i Carter przyjęli, że suma parametrów b_x dla wszystkich grup wieku (indeksowanych zmienną x) jest równa 1, a suma parametrów k_t (indeksowanych zmienną t) jest równa 0. Z ostatniego założenia wynika, że każdy z parametrów a_x w modelu (1) opisuje średni poziom umieralności w poszczególnych grupach wieku, uśredniony względem czasu kalendarzowego.

Parametry b_x wskazują, jak szybko logarytmy częściowych współczynników zgonów, tj. $\ln m_x(t)$, zmieniają w odpowiednich grupach wieku x wraz ze zmianą wartości k_t . Z kolei, k_t są parametrami opisującymi efekt wpływu czasu kalendarzowego t na zmianę częściowych współczynników zgonów.

Zauważymy, że parametry a_x i b_x w modelu (1) są stałe w czasie, co oznacza, że raz wyznaczone oceny tych parametrów mogą być ich oszacowaniami także w przyszłych okresach. Wskaźnikami zmieniającymi się wraz z upływem czasu są tu indeksy k_t , traktowane jako pewien proces stochastyczny z czasem dyskretnym. Prognozy dotyczące przewidywanych wartości tych indeksów w przyszłości, w połączeniu z oszacowaniami parametrów a_x i b_x , pozwalają, na mocy modelu (1), na sporządzenie wieloletnich prognoz umieralności, a dokładniej – prognoz dotyczących wartości współczynników zgonów w poszczególnych grupach wieku x dla przyszłych okresów.

Na podstawie przedstawionego modelu Lee i Cartera można stworzyć dynamiczne tablice trwania życia. Analizując te tablice zauważyć można istotną różnicę średniego dalszego trwania życia w zależności od wyboru tablic. Przyczyna tej różnicy jest oczywista – tablice z modelu Lee i Cartera uwzględniają proces starzenia się społeczeństwa.

Tabela 3

Tablice trwania życia z GUS-u i z modelu Lee i Cartera

Wiek	Tablice trwania życia z GUS-u		Tablice z modelu Lee i Cartera	
	mężczyźni	kobiety	mężczyźni	kobiety
60	17,69	22,94	20,56	27,43
70	11,69	15,00	14,44	19,34
80	6,95	8,35	9,03	11,92
90	3,84	4,20	5,20	6,81
99	2,21	2,22	3,77	5,03

Źródło: Dane GUS i [7].

4. Wpływ wyboru tablic trwania życia na kalkulację składki w ubezpieczeniach na życie

Wpływ ryzyka długowieczności osoby ubezpieczonej na kalkulację składki w ubezpieczeniach na życie zostanie przedstawiony na dwóch przykładach:

- renty dożywotniej dla osoby w wieku x lat, renty odroczonej o n lat (uzyskanie wieku 65 lat)
- dożywotniego ubezpieczenia na wypadek śmierci, z sumą ubezpieczenia 1 płatną na koniec roku, w którym nastąpi śmierć ubezpieczonego

W obu przypadkach dla kalkulacji składki netto przyjęte zostało $i = 4\%$.

Tabela 4

Renta dożywotnia – wysokość składki netto

Wysokość jednorazowej składki netto – renta dożywotnia				
Wiek	tablice trwania życia z GUS-u		tablice z modelu Lee i Cartera	
	mężczyźni	kobiety	mężczyźni	kobiety
20	0,0633	0,0932	0,0659	0,0954
30	0,1067	0,1544	0,1109	0,1583
40	0,1967	0,2769	0,2036	0,2833
50	0,4365	0,5900	0,4493	0,6023
60	1,7789	2,2624	1,8172	2,3039

Źródło: Opracowanie na podstawie danych GUS i [7].

Ponieważ tablice skonstruowane z modelu Lee i Cartera wykorzystują informacje z prognoz o umieralności, w przykładzie możemy zauważyć, że zarówno dla mężczyzn, jak i dla kobiet uzyskaliśmy wyższe wyniki. Wysokość jednorazowej składki netto obliczonej przy użyciu standardowych tablic trwania życia jest zdecydowanie niższa.

Tabela 5

Dożywotnie ubezpieczenie na wypadek śmierci – wysokość składki netto

Wysokość jednorazowej składki netto – dożywotnie ubezpieczenie na wypadek śmierci				
Wiek	tablice trwania życia z GUS-u		tablice z modelu Lee i Cartera	
	mężczyźni	kobiety	mężczyźni	kobiety
20	0,0072	0,0045	0,0068	0,0044
30	0,0109	0,0072	0,0104	0,0067
40	0,0172	0,0112	0,0165	0,0107
50	0,0270	0,0178	0,0258	0,0169
60	0,0426	0,0285	0,0413	0,0276

Źródło: Opracowanie na podstawie danych GUS i [7].

W przypadku dożywotniego ubezpieczenia na wypadek śmierci zarówno dla kobiet, jak i dla mężczyzn wysokość składki netto obliczonej na podstawie tablic skonstruowanych z modelu Lee i Cartera jest niższa niż składka netto obliczona za pomocą standardowych tablic trwania życia.

Podsumowanie

Przedstawione w opracowaniu wyniki pokazują, jaki wpływ ma wybór tablic trwania życia na kalkulację składki netto w ubezpieczeniach na życie. Można stąd wysnuć wniosek, że aby zmniejszyć ryzyko niewłaściwego oszacowania wartości produktów i rezerw przez zakłady ubezpieczeń na życie, należałoby rozważyć możliwość wprowadzenia do wyceny produktów ubezpieczeniowych zamiast tradycyjnych tablic trwania życia – dynamiczne tablice trwania życia. Ważne jest, aby tablice stosowane do kalkulacji składki netto wykorzystywały informacje z prognoz o umieralności, gdyż jest to dynamicznie zmieniający się w czasie proces. Zaznaczyć należy także, iż w praktyce poza umieralnością uwzględnia się jeszcze wiele innych uwarunkowań.

W literaturze przedmiotu można znaleźć uwagi sugerujące, że model Lee-Cartera można stosować do prognozowania wartości do dynamicznych tablic trwania życia tylko dla krajów, które przez wystarczająco długi okres charakteryzowały się stabilnym wzrostem oczekiwanej długości trwania życia.

Literatura

1. Lee R., Carter L. (1992). Modeling and Forecasting U.S. Mortality. *Journal of American Statistical Association*, 87 (419).
2. Oeppen J., Vaupel J. (2002). Broken Limits to Life Expectancy. *Science*, 296.
3. Olshansky J., Carnes B., Cassel C. (1990). In Search of Methuselah: Estimating the Upper Bound to Human Longevity. *Science, New Series*.
4. Papież M. (2008). Możliwość wykorzystania modelu Lee-Cartera do szacowania wartości w dynamicznych tablicach trwania życia. *Sekcja Analiz Demograficznych*, 18.
5. Rossa A. Dynamiczne tablice trwania życia oparte na metodologii Lee-Cartera i ich zastosowanie do obliczania wysokości świadczeń emerytalnych. *Acta Universitatis Lodzensis. Folia Oeconomica*, 231.
6. Sanderson W., Scherbov S. (2004). On the Road to New Stochastic Mortality Forecasts. *International Institute for Applied Systems Analysis Working Paper*.
7. Szumlicz T. (2008). Analiza ubezpieczeniowych implikacji wyników prognozy przeciętnego dalszego trwania życia uzyskanej metoda Lee i Cartera. *Wiadomości Ubezpieczeniowe – Wydanie Specjalne*.
8. Wierzchowska K. (2009). Ryzyko długowieczności w ubezpieczeniach na życie. *Zeszyty Naukowe Uniwersytet Szczeciński, Szczecin*, 550.
9. Ludność w UE: prognoza 2008-2060. Eurostat.

INFLUENCE OF RISK OF LONGEVITY ON LIFE INSURANCES

Summary

Demographic changes occurring in recent decades indicate a progressive aging of the population, and thus cause the lengthening of average life expectancy. The purpose of this article is to introduce the risk of the longevity, the methods of its prognosis and the influence on the estimation the insurance products.

Anna Nikodem-Słowikowska

Uniwersytet Ekonomiczny we Wrocławiu

ZASTOSOWANIE PROCESU MARKOWA W ZALEŻNYCH GRUPOWYCH UBEZPIECZENIACH NA ŻYCIE*

Wprowadzenie

Ubezpieczenie grupowe to ubezpieczenie polegające na objęciu ochroną większej liczby osób. Jeżeli przez T_{x_k} , $k = 1, 2, \dots, n$ oznaczmy przyszły czas trwania życia k -tej osoby w wieku x_k w grupie ubezpieczonych, to w przypadku statusu „wspólne życie”, który trwa do chwili śmierci pierwszej osoby w grupie, mamy [1; 3]

$$\begin{aligned} {}_t p_{x_1 : x_2 : \dots : x_n} &= P(\min(T_{x_1}, \dots, T_{x_n}) > t) = \\ &= P(T_{x_1} > t, \dots, T_{x_n} > t) \end{aligned} \quad (1)$$

W klasycznych ubezpieczeniach na życie zakłada się, że zmienne losowe T_{x_k} są niezależne. Zatem prawdopodobieństwo przetrwania w statusie do chwili t jest równe

$${}_t p_{x_1 : x_2 : \dots : x_n} = \prod_{k=1}^n {}_t p_{x_k} \quad (2)$$

Rozważmy zatem rentę wdowią, która po śmierci męża gwarantuje żonie wypłatę renty płaconej rocznie aż do śmierci ubezpieczonej. Jeżeli świadczenie to będzie wynosi 1 j.p., to wartość aktuarialną można wyznaczyć ze wzoru [3]

$$a_{x/y} = a_y - a_{x:y} \quad (3)$$

* Praca naukowa finansowana ze środków na naukę w latach 2010-2012 jako projekt badawczy nr 3361/B/H03/2010/38.

gdzie x oznacza wiek męża, a y wiek żony. Poszczególne składniki w składce (3) wyznacza się ze wzorów

$$a_y = \sum_{k=1}^{\omega-y} v^k p_y,$$

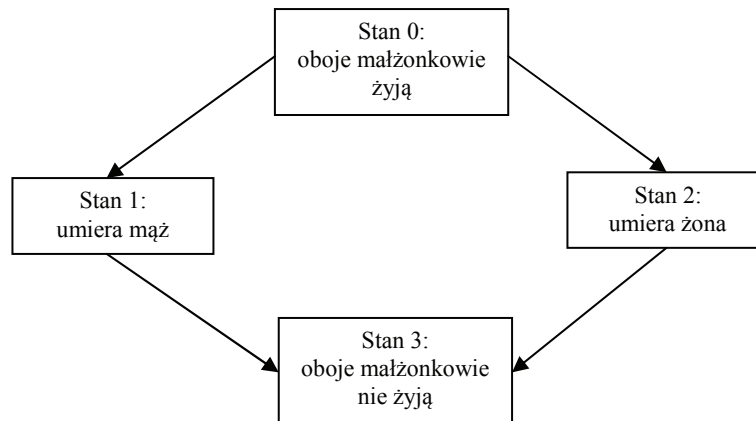
$$a_{x;y} = \sum_{k=1}^{\min\{\omega-x; \omega-y\}} v^k p_{x;y},$$

gdzie ω oznacza przyjęty maksymalny wiek, $v = (1+i)^{-1}$ jest czynnikiem dyskontującym, a i stopą procentową.

Założenie o niezależności jest jednak nierealistyczne, ponieważ np. małżeństwa narażone są na te samo ryzyko. Ponadto, prawdopodobieństwa zgonu po śmierci współmałżonka wzrasta (tzw. syndrom „złamanego serca”). Zatem wyznaczając np. składkę renty wdowiej, powinno uwzględnić się zależność między długością życia małżonków. Do zobrazowania przyszłego stanu cywilnego męża i żony, wykorzystano proces Markowa [2], który zostanie opisany niżej. Przeprowadzona zostanie także analiza wpływu zależności na wielkość składki wdowiej.

1. Zastosowanie procesu Markowa

W ubezpieczeniach dotyczących małżonków wykorzystano proces Markowa [2]. W modelu uwzględniono cztery stany, w których oboje małżonkowie żyją (stan 0), umiera jeden z małżonków (stan 1 lub 2) oraz oboje małżonkowie nie żyją (stan 3). Model ten przedstawiono na rys. 1.



Rys. 1. Model obrazujący cztery stany cywilne małżonków

Źródło: [2].

Prawdopodobieństwo przejścia w procesie Markowa przedstawionym na rys. 1 oznaczone jest przez $p_{ij}(t, t + \Delta t)$, $t, \Delta t \geq 0$ i znaczy, że małżeństwo w chwili $t + \Delta t$ jest w stanie j pod warunkiem, że w stanie i było w chwili t . Funkcja intensywności zgonów męża w wieku $x + t$, pod warunkiem, że jest on nadal żonaty, oznaczana jest przez $\mu_{01}(t)$, jeżeli jest wdowcem, to przez $\mu_{23}(t)$. Podobnie funkcja intensywności zgonów żony w wieku $y + t$ jest oznaczana przez $\mu_{02}(t)$, jeżeli jest nadal zamężna, a jeśli jest wdową, to przez $\mu_{13}(t)$. Prawdopodobieństwa przejścia dla $0 \leq t_1 \leq t_2$ można wyznaczyć za pomocą intensywności przejścia, tj.

$$p_{00}(t_1, t_2) = \exp \left\{ - \int_{t_1}^{t_2} (\mu_{01}(s) + \mu_{02}(s)) ds \right\} \quad (4)$$

$$p_{11}(t_1, t_2) = \exp \left\{ - \int_{t_1}^{t_2} \mu_{13}(s) ds \right\} \quad (5)$$

$$p_{22}(t_1, t_2) = \exp \left\{ - \int_{t_1}^{t_2} \mu_{23}(s) ds \right\} \quad (6)$$

$$p_{33}(t_1, t_2) = 1 \quad (7)$$

Dla $j = 1, 2$ prawdopodobieństwo przejścia dla $0 \leq t_1 \leq t_2$ wyznacza się ze wzoru

$$p_{0j}(t_1, t_2) = \int_{t_1}^{t_2} p_{00}(t_1, s) \mu_{0j}(s) p_{jj}(s, t_2) ds \quad (8)$$

Prawdopodobieństwo łączne wynosi [2]

$$\begin{aligned} P(T_x > t_1, T_y > t_2) = \\ = \begin{cases} p_{00}(0, t_2) + p_{00}(0, t_1) p_{01}(t_1, t_2), & \text{jeżeli } 0 \leq t_1 \leq t_2, \\ p_{00}(0, t_1) + p_{00}(0, t_2) p_{02}(t_2, t_1), & \text{jeżeli } 0 \leq t_2 \leq t_1. \end{cases} \end{aligned} \quad (9)$$

W przedstawionym modelu spełnione jest [2]

$$1) \mu_{01} \equiv \mu_{23} \text{ i } \mu_{02} \equiv \mu_{13} \Leftrightarrow T_x \text{ i } T_y \text{ są niezależne} \quad (10)$$

$$2) \mu_{01} \leq \mu_{23} \text{ i } \mu_{02} \leq \mu_{13} \Rightarrow \text{Cov}[T_x, T_y] \geq 0 \quad (11)$$

gdzie T_x i T_y oznacza odpowiednio przyszły czas życia męża i żony.

Zakładając, że intensywność zgonów osób będących w związkach małżeńskich jest niższa niż intensywność zgonów w populacji oraz że intensywność zgonów wdowców i wdów jest wyższa niż intensywność zgonów w populacji, poszczególne funkcje intensywności można zapisać jako

$$\begin{aligned} \mu_{01}(t) &= (1 - \alpha_{01})\mu_{x+t}, & \mu_{02}(t) &= (1 - \alpha_{02})\mu_{y+t} \\ \mu_{13}(t) &= (1 + \alpha_{13})\mu_{y+t}, & \mu_{23}(t) &= (1 + \alpha_{23})\mu_{x+t} \end{aligned} \quad (12)$$

gdzie parametry α_{ij} są nieujemne i α_{0j} są mniejsze od 1. W kolejnym punkcie przedstawione zostaną estymatory parametrów α_{ij} [2].

2. Estymacja parametrów

Estymatory parametrów α_{ij} minimalizują sumę kwadratów różnic między przyrostem $\Delta\Omega_{ij}$ funkcji przejścia określonej wzorem

$$\Omega_{ij}(t) = \int_0^t \mu_{ij}(s) ds \quad \text{dla } t \geq 0$$

a jego estymatorem, czyli [2]

$$\hat{\alpha}_{ij} = \arg \min \sum_k \left(\Delta \hat{\Omega}_{ij}(k) - \int_0^1 \mu_{ij}(k+t) dt \right)^2 \quad (13)$$

Estymator $\hat{\Omega}_{ij}(t)$ wyraża się wzorem [2]

$$\hat{\Omega}_{ij}(t) = \int_0^t \frac{\mathbf{1}[L_i(s) > 0]}{L_i(s)} dL_{ij}(s) \quad (14)$$

gdzie $\mathbf{1}[A]$ jest indykatozem zbioru A , $L_i(t)$ oznacza liczbę par w stanie i w wieku t , zaś $L_{ij}(t)$ jest to liczba przejść ze stanu i do stanu j do chwili t . Zakładając liniowość w ciągu roku dla całkowitych k i $0 \leq t \leq 1$ otrzymuje się

$$\begin{aligned}
L_{ij}(k+t) &= L_{ij}(k) + t \{L_{ij}(k+1) - L_{ij}(k)\} \\
L_i(k+t) &= L_i(k) + t \{L_i(k+1) - L_i(k)\}
\end{aligned} \tag{15}$$

Korzystając z (14) i (15) mamy

$$\begin{aligned}
\Delta \hat{\Omega}_{ij}(k) &= \int_0^1 \frac{\mathbf{1}[L_i(k+s) > 0]}{L_i(k) + s \{L_i(k+1) - L_i(k)\}} \{L_{ij}(k+1) - L_{ij}(k)\} ds = \\
&= \frac{L_{ij}(k+1) - L_{ij}(k)}{L_i(k+1) - L_i(k)} \{\ln L_i(k+1) - \ln L_i(k)\} \equiv \\
&\equiv \frac{L_{i,j}(k)}{L_i(k+1) - L_i(k)} \{\ln L_i(k+1) - \ln L_i(k)\}
\end{aligned} \tag{16}$$

gdzie $L_i(k)$ oznacza liczbę małżeństw w stanie i w wieku k , zaś $L_{i,j}(k)$ określa liczbę przejść ze stanu i do stanu j obserwowaną w ciągu roku. Na podstawie danych estymatory $\Delta \hat{\Omega}_{ij}(k)$ można wyznaczyć następująco:

1. Estymator $\Delta \hat{\Omega}_{01}(k)$:

$L_{0,1}(k)$ = liczba żonatych w wieku k , którzy umarli w ciągu danego roku,

$L_0(k+1) - L_0(k)$ = - liczba żonatych w wieku k , którzy umarli w ciągu danego roku - liczba żonatych w wieku k , których żona zmarła w ciągu danego roku + liczba mężczyzn w wieku k , którzy zawarli związek małżeński - liczba żonatych w wieku k , którzy wzięli rozwód w ciągu danego roku,

$L_0(k)$ = liczba żonatych na początku danego roku.

2. Estymator $\Delta \hat{\Omega}_{02}(k)$:

$L_{0,2}(k)$ = liczba mężatek w wieku k , które umarły w ciągu danego roku,

$L_0(k+1) - L_0(k)$ = - liczba mężatek w wieku k , które umarły w ciągu danego roku - liczba mężatek w wieku k , których mąż zmarł w ciągu danego roku + liczba kobiet w wieku k , które zawarły związek małżeński - liczba mężatek w wieku k , które wzięły rozwód w ciągu danego roku,

$L_0(k)$ = liczba mężatek na początku danego roku.

3. Estymator $\Delta\hat{\Omega}_{13}(k)$:

$L_{1:3}(k)$ = liczba wdów w wieku k , które umarły w ciągu danego roku,

$L_1(k+1) - L_1(k)$ = - liczba wdów w wieku k , które umarły w ciągu danego roku + liczba kobiet w wieku k , których mąż zmarł w ciągu danego roku - liczba wdów w wieku k , które zawarły związek małżeński,

$L_1(k)$ = liczba wdów na początku danego roku [2].

Intensywność zgonów według prawa Mackehama ma postać [3].

$$\mu_x = A + Bc^x \quad (17)$$

Na podstawie danych dotyczących wieku z zakresu (v_1, v_2) podzielonego na dwa zbioru $(v_1, 40)$ i $(41, v_2)$, parametry A , B i c można wyznaczyć ze wzorów [2]

$$\ln c = \frac{\sum_{x=41}^{v_2} x \cdot \sum_{x=41}^{v_2} \ln(\alpha_x) - (v_2 - 40) \sum_{x=41}^{v_2} x \ln(\alpha_x)}{\left(\sum_{x=41}^{v_2} x \right)^2 - (v_2 - 40) \sum_{x=41}^{v_2} x^2} \quad (18)$$

$$A = \frac{1}{41 - v_1} \sum_{x=v_1}^{40} (\alpha_x - \beta c^x) \quad (19)$$

$$B = \ln(c) \frac{\beta}{c - 1} \quad (20)$$

$$\text{gdzie } \alpha_x = -\ln(\hat{p}_x), \ln \beta = \frac{1}{v_2 - 40} \left(\sum_{x=41}^{v_2} \ln(\alpha_x) - \ln(c) \sum_{x=41}^{v_2} x \right)$$

Zatem korzystając z (13) i (18)-(20) estymatory $\hat{\alpha}_{ij}$ parametrów α_{ij} są równe [2]

$$\hat{\alpha}_{01} = 1 - \frac{\sum_k \left(A_1 + B_1 c_1^k \frac{c_1 - 1}{\ln c_1} \right) \Delta\hat{\Omega}_{01}(k)}{\sum_k \left(A_1 + B_1 c_1^k \frac{c_1 - 1}{\ln c_1} \right)^2} \quad (21)$$

$$\hat{\alpha}_{02} = 1 - \frac{\sum_k \left(A_2 + B_2 c_2^k \frac{c_2 - 1}{\ln c_2} \right) \Delta \hat{\Omega}_{02}(k)}{\sum_k \left(A_2 + B_2 c_2^k \frac{c_2 - 1}{\ln c_2} \right)^2} \quad (22)$$

$$\hat{\alpha}_{13} = \frac{\sum_k \left(A_2 + B_2 c_2^k \frac{c_2 - 1}{\ln c_2} \right) \Delta \hat{\Omega}_{13}(k)}{\sum_k \left(A_2 + B_2 c_2^k \frac{c_2 - 1}{\ln c_2} \right)^2} - 1 \quad (23)$$

$$\hat{\alpha}_{23} = \frac{\sum_k \left(A_1 + B_1 c_1^k \frac{c_1 - 1}{\ln c_1} \right) \Delta \hat{\Omega}_{23}(k)}{\sum_k \left(A_1 + B_1 c_1^k \frac{c_1 - 1}{\ln c_1} \right)^2} - 1 \quad (24)$$

3. Składka renty wdowiej w modelu Markowa

Składka renty wdowiej w przypadku zastosowania procesu Markowa ma postać [2]

$$a_{x|y} = \sum_{k=0}^{\min\{\omega_x - x; \omega_y - y\}} p_{00}(0, k) p_{01}(k, k+1) \cdot \sum_{j=0}^{\omega_y - y - k} p_{11}(k+1, k+1+j) v^{k+1+j}, \quad (25)$$

gdzie ω_x, ω_y oznaczają maksymalny wiek. Stosując (4)-(8) oraz (12) i (17) poszczególne prawdopodobieństwa we wzorze na składkę wynoszą

$$p_{00}(0, k) = \exp \left\{ -(1 - \alpha_{01}) \left(A_1 \cdot k + \frac{B_1}{\ln c_1} (c_1^{x+k} - c_1^x) \right) + \right. \\ \left. -(1 - \alpha_{02}) \left(A_2 \cdot k + \frac{B_2}{\ln c_2} (c_2^{y+k} - c_2^y) \right) \right\},$$

$$p_{11}(k, k+1+j) = \exp \left\{ -(1 - \alpha_{13}) \left(A_2 \cdot j + \frac{B_2}{\ln c_2} (c_2^{y+k+1+j} - c_2^{y+k+1}) \right) \right\}$$

oraz przy założeniu, że intensywności przejścia w ciągu roku są stałe

$$p_{01}(k, k+1) = \frac{\mu_{01}(k) (\exp\{-\mu_{01}(k) - \mu_{02}(k)\} - \exp\{-\mu_{13}(k)\})}{\mu_{13}(k) - \mu_{01}(k) - \mu_{02}(k)}$$

4. Przykład

Na podstawie danych demograficznych z 2002 roku (dane pochodzą ze strony internetowej GUSu [4]) przedstawione zostaną wartości estymatorów oraz składki renty wdowiej korzystając z procesu Markowa oraz z założenia o niezależności.

Wielkości oszacowanych parametrów dla funkcji intensywności zgonów dla żony (nr 2) i męża (nr 1) przedstawiono w tabeli 1. W tabeli 2 umieszczono wartości estymatorów parametrów α_{ij} .

Tabela 1

Wartość estymatorów funkcji intensywności zgonów dla męża i żony

i	1	2
s_i	0,999944169	0,999783701
g_i	0,997987825	0,999748444
c_i	1,082979413	1,101529514
A_i	0,000055833	0,000216323
B_i	0,000160564	0,000024328

Tabela 2

Wartości estymatorów parametrów α_{ij}

$\hat{\alpha}_{01}$	0,2317634
$\hat{\alpha}_{02}$	0,1827206
$\hat{\alpha}_{13}$	0,2106567

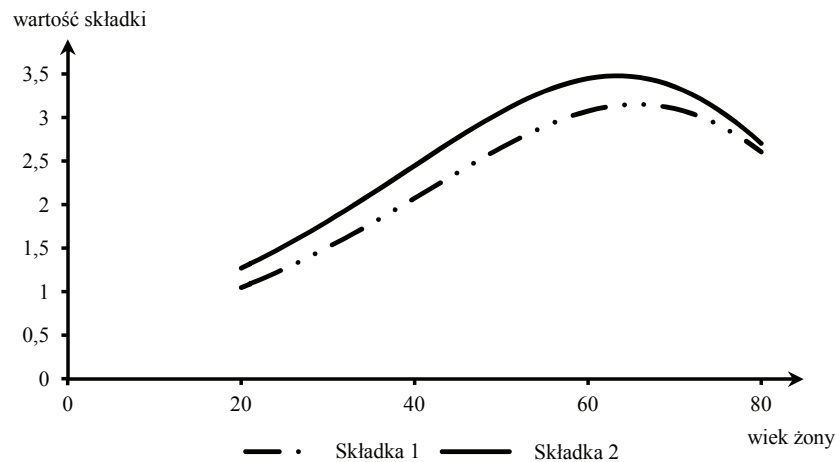
W zależności od wieku męża (x) i wieku żony (y) wyznaczono składkę renty wdowiej w modelu Markowa jak i przy założeniu niezależności. Przyjęto stopę techniczną równą 5%. W ostatniej kolumnie tabeli 3 obliczono, jaki procent składka w modelu Markowa stanowi składkę w modelu zakładającym niezależność.

Tabela 3

Wielkość składki w zależności o wieku męża i żony

Wiek żony y	Wiek męża x	Składka 1 w modelu Markowa	Składka 2 przy założeniu niezależności	Wielkość składki 1 w porównaniu ze składką 2 (%)
25	27	1,432332	1,723144	83,12
32	28	1,254259	1,50802	83,17
40	42	2,336496	2,748325	85,02
52	55	3,263514	3,725365	87,60
60	61	3,247426	3,633255	89,38

Na wykresie na rys. 2 przedstawiono wartość składki w modelu Markowa oraz wartość składki przy założeniu niezależności dla małżeństw, w których mąż i żona są w tym samym wieku. Do obliczenia składek przyjęto stopę procentową i równą 5%.



Rys. 2. Wartość składki w modelu Markowa i w przypadku niezależności

Na podstawie wyników z tabeli 3 widać, że składka, zakładając zależność opisaną procesem Markowa, jest niższa od składki wyliczonej przy założeniu niezależności. Ponadto można zauważyć także, że w przypadku młodszych ubezpieczonych ta różnica względna jest większa, niż w przypadku starszych ubezpieczonych. Można to też zaobserwować na wykresie przedstawionym na rys. 2.

Literatura

1. Błaszczyszyn B., Rolski T. (2004). Podstawy matematyki ubezpieczeń na życie. Wydawnictwo Naukowo-Techniczne, Warszawa.
2. Denuit M., Dhaene J., Tillegheem L., Teghem S. (2001). Measuring the Impact of a Dependence Among Insured Lifespans. *Belgian Actuarial Bulletin*, 1, 1, s. 18-39.
3. Skarba M. (1999). Ubezpieczenia na życie. Wydawnictwo Naukowo-Techniczne, Warszawa.
4. Witryna internetowa Głównego Urzędu Statystycznego: www.stat.gov.pl

APPLICATION OF A MARKOV PROCESS IN THE MULTIPLE LIFE INSURANCE WITH DEPENDENT LIFE LENGTHS

Summary

Consider the multiple life insurance. Let T_k denote the remaining lifetimes of a k th person in the group. In the classical theory of multiple life insurance it is assumed that the random variables T_k are independent. This assumption is unrealistic. For example, considering the lifetimes of paired lives such as wife and husband one should take into consideration the dependence of the lifetimes of this married couple.

In this article, the future development of the marital status for husband and wife will be represented as a Markov process. On the basis of the data the parameters of this model will be estimated. The influence on the amount of premium relating to the widow's pension will also be studied.

Stanisław Wieteska

Uniwersytet Łódzki

UBEZPIECZENIE LASÓW OD POŻARÓW W POLSKIEJ STREFIE KLIMATYCZNEJ

Wprowadzenie

Jednym z zagrożeń dla lasów w Polsce są pożary. Zagrożenie pożarami było wielokrotnie przedmiotem wielu opracowań. Pożary lasów to straty w drzewostanie, ekologii terenów leśnych, w faunie zwierzęcej. Pomimo ogromnego postępu w zakresie zwalczania i prewencji pożarowej, w dalszym ciągu występują one na terenie Polski, powodując ogromne straty.

Warto jednak zastanowić się nad możliwością rekompensowania strat pożarowych przez zakłady ubezpieczeń majątkowo-osobowych. Ubezpieczenie lasów od pożarów należy traktować jako jeden z instrumentów finansowych, chroniący majątek narodowy. Obecnie obserwujemy zjawisko zmniejszających się środków finansowych (środków publicznych) na utrzymanie lasów [15, s. 6-7]. Stąd potrzeba ubezpieczenia lasów od pożarów. Ubezpieczenie to ma także za zadanie chronić gospodarstwa leśne przed stratami spowodowanymi przez siły przyrody i człowieka. Ubezpieczenie to także zapewnienie stabilności w produkcji leśnej.

Celem tego opracowania jest więc próba prezentacji podstaw do kalkulacji stopy składki dla strat spowodowanych pożarami lasów.

Analizę pożarów lasów ograniczamy do polskiej strefy klimatycznej. Przez strefę klimatyczną rozumiemy będziemy charakterystyczny dla Polski zespół zjawisk i procesów atmosferycznych kształtowanych pod wpływem uwarunkowań fizycznych i geograficznych określonych na podstawie wieloletnich obserwacji.

Opracowanie napisano opierając się na załączonej literaturze przedmiotu, danych statystycznych GUS i Leśnictwa. Rozważa się tylko te podstawowe elementy, które są niezbędne do celów kalkulacji stóp składek.

1. Klasyfikacja i przyczyny pożarów lasów

W literaturze spotykamy następującą klasyfikację pożarów leśnych:

- ziemny (podziemny), gdy palą się pokłady torfu, węgla brunatnego; występują najczęściej pod koniec lata i po długotrwałych suszach, trwać mogą kilka miesięcy, z ziemi wydobywa się gryzący dym, rozprzestrzeniający się stosunkowo wolno; są niebezpieczne, gdyż pod powierzchnią tworzą się zapadliny trudne do zauważenia,
- przyziemne, gdy pali się runo leśne, podszycie, zewnętrzna kora oraz płytko zalegające korzenie; mogą być przelotne lub trwałe, szybkość rozprzestrzeniania się tych pożarów uzależniona jest od materiału palnego,
- wierzchołkowy, czyli górny, przesuwany się po konarach drzew, spala się igliwie, osmalają się gałęzie i pnie drzew [2, s. 56-60].

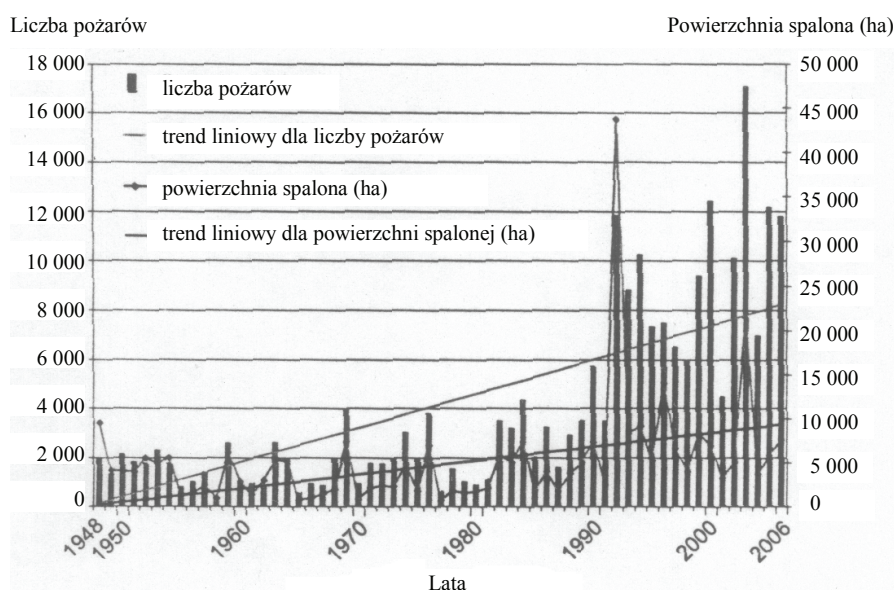
Każdy rodzaj pożaru lasu spowodowany jest przez różne przyczyny. Do najważniejszych przyczyn należy zaliczyć według klasyfikacji GUS:

- nieostrożne obchodzenie się z ogniem przez dzieci i dorosłych,
- podpalenia,
- wyładowania atmosferyczne,
- wady urządzeń technicznych i nieprawidłowa ich eksploatacja,
- wady środków transportowych i ich nieprawidłowa eksploatacja (koleje, ciągniki, pojazdy samochodowe),
- pozostałe,
- nieustalone.

Z punktu widzenia ubezpieczeniowego ochroną ubezpieczeniową powinniśmy obejmować pożary powstałe w sposób losowy. Do przypadków losowych możemy zaliczyć: wyładowania atmosferyczne oraz wady urządzeń technicznych i transportowych wraz z ich nieprawidłową eksploatacją.

Przebieg pożarów w lasach uzależniony jest od wielu czynników. Pogoda jest w dalszym ciągu najtrudniejszym do przewidzenia elementem mającym wpływ na przebieg pożaru [10, s. 152-156]. Dlatego też konieczne jest opracowywanie prognoz odnośnie do rozwoju pożarów w różnych warunkach atmosferycznych.

Warto także zwrócić uwagę, że liczba pożarów lasów uzależniona jest od zmian klimatycznych [14, s. 67-71]. Badania wieloletnich ciągów danych meteorologicznych doprowadziły do wniosku, że rosnący trend temperatury powietrza oraz malejąca tendencja opadów atmosferycznych od początku XX wieku w Polsce przyczyniły się do wzrostu zagrożenia pożarowego lasów (rys. 1).



Rys. 1. Liczba pożarów lasów i powierzchnia spalona w Polsce w latach 1948-2006 oraz charakteryzujące je trendy

Źródło: [14, s. 70].

W wyniku zmian ocieplenia klimatu nastąpił wzrost liczby pożarów i powierzchni spalonej lasów, szczególnie widoczny w latach dziewięćdziesiątych XX wieku. Należy podkreślić, że trend wzrostu liczby pożarów był większy niż trend wzrostu powierzchni spalonej przez pożary. Świadczyć to może o systematycznej poprawie w zakresie organizacji likwidacji pożarów i monitoringu ochrony przeciwpożarowej. W związku z przewidywanymi scenariuszami dotyczącymi dalszych zmian klimatu, należy się liczyć z nasileniem zjawisk pożarów. Dla potrzeb (prognozowania) przewidywania kierunku rozwoju pożarów, straże pożarne powinny wziąć pod uwagę między innymi rodzaj materiału palnego, ukształtowanie terenu, lokalizację zbiorników wodnych, rzek, przewidywanie możliwej wielkości obszaru objętego pożarem itp.

Możemy powiedzieć, że przebieg pożaru ma charakter losowy i trudny do przewidzenia.

2. Analiza statystyczna pożarów w Polsce w latach 1999-2008

Ewidencję statystyczną pożarów prowadzi od wielu lat Straż Pożarna a także Główny Urząd Statystyczny. Liczba pożarów lasów w ciągu roku jest zmienna (tabela 1).

Tabela 1

Liczba pożarów lasów według wybranych przyczyn powstania w Polsce w latach 1999-2008

Rok	Ogółem	Przyczyny powstania		
		wyladowania atmosferyczne	wady urządzeń technicznych i ich nieprawidłowa eksploatacja	wady środków transportu i ich nieprawidłowa eksploatacja
1999	9405	26	55	12
2000	12428	64	61	29
2001	4480	15	21	3
2002	10101	41	45	7
2003	17088	67	69	20
2004	7006	19	29	5
2005	12169	33	40	13
2006	11828	65	71	32
2007	7101	51	23	11
2008	9091	37	89	70
Średnio	10069,7	41,8	50,3	20,2
Struktura procentowa	1,00	0,0041	0,0050	0,0020

Źródło: [7].

Z danych zawartych w tabeli 1 wynika, że ok. 2‰-5‰ pożarów spowodowana jest przez wyladowania atmosferyczne i wady urządzeń technicznych i transportowych. Analogiczna sytuacja występuje w przypadku powierzchni pożarów (tabela 2).

Tabela 2

Powierzchnia pożarów lasów (w ha) według wybranych przyczyn powstania w Polsce w latach 1999-2008

Rok	Ogółem	Przyczyny powstania		
		wyladowania atmosferyczne	wady urządzeń technicznych i ich nieprawidłowa eksploatacja	wady środków transportu i ich nieprawidłowa eksploatacja
1999	8307	8	102	13
2000	7013	13	15	21
2001	3429	6	8	3
2002	5200	7	26	18
2003	21500	10	80	22
2004	3781	2	23	7
2005	5826	4	22	25
2006	5912	6	72	31
2007	3564	16	3	9
2008	3028	3	53	27
Średnio	6756,0	7,5	40,4	17,6
Struktura procentowa	1,00	0,0011	0,0060	0,0026

Źródło: [7].

Z danych zawartych w tabeli 2 wynika, że najmniejsza powierzchnia pożarów jest spowodowana wyladowaniami atmosferycznymi. Widzimy wyraźnie regionalne zróżnicowanie liczby pożarów w poszczególnych województwach (tabela 3).

Tabela 3

Regionalne zróżnicowanie liczby pożarów w lasach w Polsce w latach 2003-2004 i 2008

Lp.	Województwa	2003	2004	2005	2008
1	2	3	4	5	6
1	razem	156	53	86	196
2	dolnośląskie	5	2	–	10
3	kujawsko-pomorskie	7	3	8	10
4	lubelskie	7	2	5	2
5	lubuskie	27	9	13	46
6	łódzkie	13	7	13	10
7	małopolskie	1	1	–	3

cd. tabeli 3

1	2	3	4	5	6
8	mazowieckie	21	2	16	15
9	opolskie	5	2	2	4
10	podkarpackie	7	1	–	1
11	podlaskie	4	3	2	4
12	pomorskie	4	1	2	2
13	śląskie	10	1	9	9
14	świętokrzyskie	5	2	–	5
15	warmińsko-mazurskie	2	1	1	10
16	wielkopolskie	33	13	12	49
17	zachodniopomorskie	5	3	3	16

Uwaga: Pożary spowodowane przez: wyladowania atmosferyczne, wady urządzeń technicznych i transportowych wraz z ich niewłaściwą eksploatacją.

Źródło: [7].

Z danych zawartych w tabeli 3 wynika, że najczęściej pożarów powstaje w województwach: lubuskie, łódzkie, mazowieckie, śląskie, wielkopolskie.

3. Kategoryzacja zagrożenia pożarowego obszarów leśnych

Obowiązek klasyfikacji lasów pod względem zagrożenia pożarowego został nałożony na kraje członkowskie Unii Europejskiej Rozporządzeniem Rady (E E C) nr 2/58/92 z dnia 23 lipca 1992 roku o ochronie lasów Wspólnoty przed pożarami. Artykuł 2 Rozporządzenia ustalił, że kraje członkowskie klasyfikują swoje terytoria według stopnia ryzyka pożarowego w zależności od cykliczności ich powstawania*.

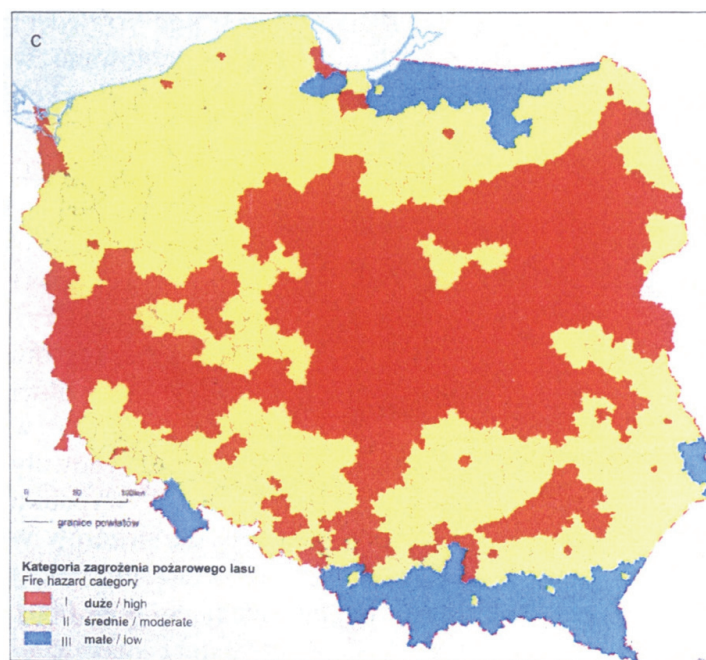
Dla celów oceny ryzyka ubezpieczeniowego w przypadku lasów warto wykorzystać ustalenia w zakresie kategoryzacji zagrożenia pożarowego.

Pod pojęciem kategorii zagrożenia pożarowego rozumieć będziemy „cechę umowną” mającą na celu nadanie obszarowi leśnemu wyróżnika cyfrowego i wyrazowego określającego jego podatność na możliwość powstania i rozprzestrzenienia się pożarów [17, s. 8].

Badania nad kryteriami zagrożenia pożarowego w lasach są prowadzone od dość dawna w wielu krajach [4, s. 3-11]. Wypracowano wiele kryteriów oceny zagrożenia pożarowego. W Polsce zasady kategoryzacji lasów określa Rozporządzenie Ministra Środowiska z dnia 22 marca 2006 roku w sprawie szczegółowych zasad ubezpieczenia przeciwpożarowego lasów (Dz.U. 2006,

* Kategoryzacją zagrożenia pożarowego w Polsce zajmowało się wielu autorów [13, s. 131-141].

nr 58, poz. 105). Obowiązująca metoda kategoryzacji dotyczy tylko nadleśnictw i parków narodowych. Uznano, że kategoryzacja powinna dotyczyć wszystkich lasów bez względu na formę własności. Opracowano kilka map zagrożenia pożarowego, biorąc pod uwagę cztery parametry, tj. występowanie pożarów, czynniki drzewostanowe, czynniki klimatyczne oraz czynniki antropogeniczne [13, s. 131-141]. Dla naszych celów prezentujemy mapę zagrożenia opartą na aktualnej bazie administracyjnej powiatów (rys. 2).



Rys. 2. Kategoryzacja zagrożenia pożarowego (31.08.2008) według powiatów (mapy sporządził R. Hildebrandt)

Źródło: [13, s. 138].

Z danych zawartych na tej mapie wynika regionalne zróżnicowanie zagrożenia pożarowego w Polsce. Widzimy wyraźnie, że każda kategoria zagrożenia powinna posiadać inną stopę składki.

W 2010 roku wydano Rozporządzenie Ministra Środowiska zmieniające Rozporządzenie w sprawie szczegółowych zasad zabezpieczania przeciwpożarowego lasów. W Rozporządzeniu podano sposób zaliczania lasów do kategorii zagrożenia pożarowego lasów. Zaliczania lasów do kategorii zagrożenia pożarowego lasu dokonuje się dla obszaru każdego nadleśnictwa albo parku narodowego o uproszczonych planach urządzenia lasu albo planach ochrony parku narodowego, na podstawie sumy punktów odpowiadających:

1) średniej rocznej liczbie pożarów lasu w okresie ostatnich 10 lat przypadających na 10 km² powierzchni leśnej,

2) udziałowi procentowemu powierzchni drzewostanów rosnących na siedliskach boru suchego, boru świeżego, boru mieszanego świeżego, boru wilgotnego, boru mieszanego wilgotnego i lasu łęgowego,

3) średniej wilgotności względnej powietrza (pomiar z wysokości 0,5 m) i procentowego udziału dni z wilgotnością ściółki mniejszą od 15% o godzinie 9⁰⁰,

4) średniej liczbie mieszkańców przypadających na 0,01 km² powierzchni leśnej.

Opracowane zostały wzory na obliczenie poszczególnych składników sumy punktów.

4. Szacowanie szkód spowodowanych przez pożary

Ważnym elementem w zakresie ubezpieczeń jest oszacowanie wartościowe strat spowodowane przez wypadek losowy. W przypadku oszacowania strat spowodowanych przez pożar w lasach mamy do czynienia z wyjątkowo trudną materią, gdyż o stratach decyduje wiele czynników.

Dyskusja nad metodami szacowania szkód pożarowych trwa już dość długo [5, s. 17-18].

W artykule S. Zająca i Janiny Parzuchowskiej pt. *Metody wyceny szkód powstałych w wyniku pożarów lasu*, autorzy dokonali przeglądu sposobów oceny strat w wybranych krajach. Autorzy zaproponowali także wzory na obliczenie strat z tytułu:

- przedwczesnego wyrębu drzewostanu częściowo i całkowicie zniszczonego,
- dodatkowych kosztów odnowienia w powierzchni drzewostanów,
- dodatkowych kosztów pozyskania w drzewostanach,
- całkowitego zniszczenia gleby,
- szkód w drzewostanach na obrzeżach pożarowych.

W dalszej kolejności zastosowali opracowane metody do oszacowania szkód pożarowych w Polsce w latach 1991-1995.

Szacunkowych danych o stratach w wyniku pożarów dostarczają nam dane GUS (tabela 4).

Tabela 4

Analiza pożarów lasów państwowych w Polsce w latach 1999-2005

Lp.	Wyszczególnienie	1999	2000	2001	2002	2003	2004	2005
1	2	3	4	5	6	7	8	9
1	Powierzchnia lasów dotkniętych pożarami w ha	1800	1766	685	1180	4182	998	1197

cd. tabeli 4

1	2	3	4	5	6	7	8	9
	– w tym drzewostany III klasy wieku i wyższych	449	616	149	400	1392	277	407
2	Przeciętna powierzchnia jednego pożaru w ha	0,4	0,35	0,33	0,31	0,51	0,29	0,27
3	Wartość strat sumy bieżące w mln zł	10,4	8,6	2,9	4,4	15,7	3,5	3,5
4	Średnia szacunkowa strata na jeden pożar w zł	5778	4870	4234	3729	3754	3507	2923

Źródło: [7, s. 305].

Z danych zawartych w tabeli 4 widzimy, że maleje procentowa powierzchnia spalonego lasu. Szacunkowy koszt strat (w zł) waha się w granicach 3000-6000 zł z tendencją malejącą.

W innym badaniu analiza ekonomiczna funkcjonowania ochrony przeciwpożarowej w lasach państwowych [13, s. 27-50] w latach 2000-2004 wykazała, że:

- średnia wartość spalonego 1 ha lasów wyniosła ok. 7441 zł (minimalna 4391, a maksymalna 11 302 zł),
- średnia roczna strata w przeliczeniu na 1 ha powierzchni zalesionej wyniosła ok. 1,02 zł (największa RDLP Katowice 2,49 najmniejsza w RDLP Krośnie 0,21 zł); biorąc pod uwagę szkody ekologiczne, średnia strata roczna odniesiona do 1 ha powierzchni zalesionej wyniosłaby ok. 5,77 zł,
- średnie roczne koszty ochrony przed pożarami wyniosły ok. 8,47 zł na 1 ha (największa w RDLP Warszawa 20,22 zł, a najniższe w Krośnie ok. 1,97 zł/ha).

W latach 2004-2008 wydano na ochronę przeciwpożarową lasów łącznie 309,1 mln zł [6, s. 274]. Roczne straty nadzwyczajne w tym czasie wynosiły 0,5-1,5 mln zł.

5. Częstość szkód pożarowych

Do oceny ryzyka pożarów w lasach w analizach używa się wskaźników:

- gęstość pożarów tzn. ilość pożarów na 10 tys. ha,
- średnia roczna spalona powierzchnia na 10 tys. ha lasów,
- średnia powierzchnia spalona podczas jednego pożaru,
- liczba punktów obserwacyjnych (dostrzegalność) na 10 tys. ha.

Dla potrzeb kalkulacji stopy składki – jak się wydaje – najlepszym wskaźnikiem częstości byłyby średnia roczna powierzchnia spalona na 10 tys. ha. Za tym wskaźnikiem przemawiają następujące argumenty:

- jest to ubezpieczenie (do 1 roku) majątkowo-osobowe,
- w pożarach płonie część lub całość ubezpieczonego mienia (lasu).

Dyskusyjna jest powierzchnia odniesienia, tj. 10 tys. ha (czyli 100 km² lasów).

Przeprowadzone badania ankietą wśród 24 państw w latach 1977-1979 na temat występowania pożarów wykazały, że z powodu wyładowań atmosferycznych a więc zdarzeń losowych powstaje ok. 2% pożarów w lasach [4, s. 10].

W latach 1991-2000 w ROLP w Zielonej Górze zanotowano 121 pożarów o powierzchni ok. 9 ha [11, s. 27]. Są to jak widać małe powierzchniowo pożary spowodowane wyładowaniami atmosferycznymi. Na podstawie tabeli 4 możemy proponować stopę składki na poziomie 2‰ wartości ubezpieczonego lasu.

6. Metody szacowania wartości drzewostanu leśnego.

Suma ubezpieczenia

Bardzo ważnym elementem umowy jest wartość przedmiotu ubezpieczenia. Wartość tę nazywa się sumą ubezpieczenia. Przez sumę ubezpieczenia rozumiemy kwotę wyrażoną wartościowo na którą ubezpieczymy drzewostan leśny. Ustalenie sumy ubezpieczenia jest sprawą złożoną ze względu na specyfikę przedmiotu ubezpieczenia. W literaturze przedmiotu podejmowane były próby wyceny wartości drzewostanu. W tym celu powinniśmy wykorzystać badania Wojciecha Wilkowskiego poświęcone metodzie określania wartości drzewostanów [16, s. 3-9]. Opracował on kilka wzorów, dla przykładu wartość drzewa określono wzorem

$$W_D = V_a \cdot r_i \left(\sum k_i \cdot U_i \cdot C_i + V_{DR} \cdot k_{DR} \cdot U_{DR} \cdot C_{DR} \right)$$

gdzie:

V_a – miąższość grubizny drzewa,

r_i – współczynnik do obliczania miąższości netto zawarty w tabeli 5,

k_i – wskaźnik udziału w grubiznie całkowitej określonego i-tego sortymentu,

C_i – cena sprzedaży i-tego sortymentu określona przez nadleśnictwo bez podatku VAT,

U_i – wskaźnik redukujący ceny sprzedaży C_i tego sortymentu o koszty pozyskania i zrywki,

V_{DRi} – miąższość drobnicy,

U_{DRi} – wskaźnik redukujący ceny sprzedaży C_{DRi} o koszty pozyskanie i zrywki,

k_{DRi} – udział procentowy drobnicy mający wartość rynkową w drobnicy całkowitej,

C_{DRi} – cena sprzedaży drobnicy przez właściwe nadleśnictwo.

Tabela 5

Współczynniki służące do obliczania miąższości netto

Rodzaj drzewa	Klasy wieku			
	II	III	IV	V i wyżej, KO, KDO, BP, przestoje
Sosna, modrzew	0,72	0,75	0,77	0,79
Świerk, jodła, daglezja	0,78	0,80	0,82	0,82
Dąb, jesion, jawor, wiąz, klon i inne twardo liściaste	0,70	0,72	0,73	0,78
Buk, grab	0,81	0,84	0,86	0,87
Osika, topola, wierzba	0,74	0,71	0,79	0,80
Brzoza, olcha	0,73	0,75	0,77	0,79

Źródło: Instrukcja urządzenia lasu, część ogólna. Ministerstwo Ochrony Środowiska, Zasobów Naturalnych i Leśnictwa, Dyrekcja Lasów Państwowych, Warszawa 1994.

W odniesieniu do drzewostanów stanowiących drągowiznę stosuje się wzór uproszczony

$$W = W_{sp} \cdot Z \cdot P \cdot C \cdot U$$

gdzie:

W_{sp} – przeliczniki wyrażone w m³ surowca tartacznego (iglastego WBO) dla określonego gatunku, klasy bonitacji, wieku (wiek rębności określono na podstawie tablic opracowanych w Instytucie Badawczym Leśnictwa, autorzy: T. Partyka, T. Trampler)*,

Z – stopień zadrzewienie,

P – powierzchnia zajmowana przez drzewostan,

C – cena sprzedaży m³ drewna tartacznego iglastego WBO ustalona przez to nadleśnictwo,

U – wskaźnik redukujący ceny sprzedaży drewna o koszty pozyskania i zrywki.

* Przeliczniki WSP zawarte są w Zarządzeniu Ministra Ochrony Środowiska Zasobów Naturalnych i Leśnictwa z 30.12.1995 (MP – 3 z 1946 r.) zał. 1.

W drzewostanach mieszanych ich wartość stanowi suma

$$\tilde{W} = \sum_{i=1}^n (W_{spi} \cdot Z_i) \cdot C \cdot P \cdot U$$

gdzie:

- W_{spi} – przelicznik w m² surowca tartacznego iglastego WBO dla i -tego gatunku, uzyskany z określonych tablic,
- Z_i – częstkowy stopień zadrzewienia i -tego gatunku.
- C – cena sprzedaży m³ drewna tartacznego iglastego WBO ustalona przez to nadleśnictwo,
- P – powierzchnia zajmowana przez drzewostan,
- U – wskaźnik redukujący ceny sprzedaży drewna o koszty pozyskania i zrywki.

W warunkach gospodarki rynkowej o wartości drzewostanu decyduje cena rynkowa oraz cena asortymentu drzewa. Cena asortymentu obejmuje koszty: pozyskania, wywozu i sprzedaży drzewa.

Dla celów ustalenia sumy ubezpieczenia koniecznością byłoby opracowanie szacunkowych cen 1 ha lasów w danym roku przez praktyków zajmujących się leśnictwem w Polsce. Obliczenia te wymagają przeprowadzenia inwentaryzacji obszarów leśnych, oceny kondycji drzew oraz opracowania leśnej mapy numerycznej o wartości drzewostanu leśnego.

Według danych statystycznych GUS tabela 4 za 2008 rok, cena 1m³ drzewa leśnego wyniosła 148,95 zł, zasoby zaś 1 ha lasu wynoszą 240 m³ grubizny drzewnej. Mnożąc te wielkości przez siebie uzyskujemy cenę 1 ha lasów na ok. 36 tys. zł.

Warto w tym miejscu podkreślić, że suma ubezpieczenia jest górną granicą odpowiedzialności zakładu ubezpieczeń. Powinna być wielkością wyczerpywalną. Wypłata odszkodowania powinna być pomniejszona o wartość sprzedaży drzewa popożarowego.

7. Proponowany zakres ubezpieczenia lasów od pożarów

Podstawą ubezpieczenia są ogólne warunki ubezpieczeń. Treść ogólnych warunków ubezpieczeń określa art. 12a ustawy o działalności ubezpieczeniowej z 22 maja 2003 r. (załącznik do obwieszczenia Marszałka Sejmu Rzeczypospolitej Polski z dnia 16 grudnia 2009 r., poz. 66).

Proponowanym ubezpieczeniem od ognia można by objąć obszary leśne będące pod zarządem państwowym, prywatnym lub innych organizacji np. samorządu terytorialnego.

Ochrona ubezpieczeniowa obejmowałaby tylko przypadki powstania pożarów na skutek wyładowań elektrycznych, zwarć w liniach wysokiego napięcia, wady środków transportu kolejowego i ich nieprawidłowa eksploatacja, wady urządzeń technicznych, energetycznych i ich nieprawidłowa eksploatacja. Każda z przyczyn pożarów powinna być udowodniona i udokumentowana.

Ubezpieczeniem objęty powinien być ściśle określony obszar leśny, w ramach granic wyznaczonych przez mapy geodezyjne.

Ubezpieczenia nie obejmowały problemów regeneracji (pożarzysk) ekologicznych. W ramach ubezpieczenia za dodatkową składkę może ono pokrywać część lub całość szkód związanych z akcją likwidacji bezpośredniej pożarów lasów.

Ubezpieczony obszar leśny powinien zawierać odnogi dojazdowe, stały system wykrywania pożarów (punkty stałe naziemnej obserwacji), pasy przeciwpożarowe, bazy sprzętu przeciwpożarowego, źródła wody do celów pożarowych, oraz oznakowania wstępu do lasów. Ponadto spełnione powinny być przepisy bezpieczeństwa pożarowego, środki łączności, systemy szkolenia [18, s. 8-9; 9, s. 12-13], aktualne mapy ochrony przeciwpożarowej [8, s. 15-16].

Warto także w ogólnych warunkach ubezpieczeń wskazać na konieczność wyposażenia obszarów leśnych w nowe automatyczne stacje prognozy meteorologicznej [1, s. 18-19]. Pozwolić to może na prowadzenie skutecznego monitoringu pożarowego w lasach na terenie całego kraju.

Dla potrzeb ubezpieczeniowych, zwłaszcza przy likwidacji szkód spowodowanych przez pożary, należy wziąć pod uwagę interwencjonizm państwa w likwidacji szkód spowodowanych przez żywioły. Interwencjonizm państwa odbywa się za pomocą między innymi aktów prawnych. Są to:

1. Rozporządzenie Ministra Rolnictwa i Rozwoju Wsi z dnia 13 lipca 2010 r. poz. 870 zmieniające rozporządzenie w sprawie szczegółowych warunków i trybu przyznawania pomocy finansowej w ramach działania dotyczącego odtwarzania potencjału produkcji leśnej zniszczonej przez katastrofy oraz wprowadzenia instrumentów zapobiegawczych objętego programem Rozwoju Obszarów Wiejskich na lata 2007-2013 (Dz.U. 2010, poz. 870). W rozporządzeniu tym jest wymieniona lista nadleśnictw uprawnionych do ubiegania się o taką pomoc.

2. Rozporządzenie Ministra Środowiska z 12.III.2006 r. w sprawie szczegółowych zasad zabezpieczenia przeciwpożarowego lasów (Dz.U. 2006, nr 58, poz. 405).

Ustalenia powyższych rozporządzeń powinny być włączone do ogólnych warunków ubezpieczeń proponowanego ubezpieczenia.

Podsumowanie

Ubezpieczenie lasów od pożarów jest nowym obszarem do objęcia go ochroną ubezpieczeń majątkowych. Ubezpieczenie to może funkcjonować w ramach ubezpieczenia od ognia i innych zdarzeń losowych. Jak się okazuje, istnieje w Polsce dość dobrze prowadzona statystyka pożarów w lasach, co pozwala na kalkulacje stóp składek. Podejmowane są próby oszacowania wartościowego strat pożarów w lasach oraz kategoryzacji zagrożeń pożarowych. Ochroną ubezpieczeniową powinny być objęte pożary lasów z przyczyn losowych, do których zaliczylibyśmy: wyładowania atmosferyczne i wady urządzeń technicznych i pojazdów samochodowych, kolejowych.

Praca nie wyczerpała problematyki, lecz ją zasygnalizowała. Konieczne są ustalenia w zakresie nadleśnictw w Polsce.

Literatura

1. Automatyczne prognozy pożarowe. (2008). Las Polski, 13-14.
2. Karbowski W. (1972). Prognozowanie i klasyfikacja pożarów leśnych. Biuletyn Informacji Technicznej, 3.
3. Karlikowski T., Szczygieł R. (1972). Kryteria oceny i prognozowania zagrożenia pożarowego lasów w wybranych krajach. Biuletyn Informacji Technicznej, 3.
4. Karlikowski T., Szczygieł R. (1982). Występowanie pożarów leśnych. Biuletyn Informacji Technicznej, 1-2.
5. Kietliński B. (1958). O trudnościach szacowania szkód spowodowanych pożarami leśnymi. Las Polski, 13-14.
6. Leśnictwo. (2009). GUS, Warszawa.
7. Leśnictwo. (2004-2006). GUS, Warszawa.
8. Mycke-Dominika M. (2002). Mapa ochrony przeciwpożarowej lasu. Las Polski, 6.
9. Niemiec P. (2003). Zintegrowany system przeciwpożarowy. Las Polski, 4.
10. Perry D.G. (1982). Model pożaru leśnego pomaga w przewidywaniu jego przebiegu oraz w przygotowywaniu planów obrony. Biuletyn Informacji Technicznej, 1-2.
11. Skowron K. (2003). Wyładowania atmosferyczne a pożary lasów. Las Polski, 4.
12. Szczygieł R., Piwnicki J., Ubysz B. (2007). Analiza ekonomiczna funkcjonowania ochrony przeciwpożarowej lasu w lasach państwowych. Leśne Prace Badawcze, 1.
13. Szczygieł R., Ubysz B. i inni. (2009). Klasyfikacja zagrożenia pożarowego lasów Polski. Leśne Prace Badawcze, 70 (2).
14. Szczygieł R., Ubysz B., Piwnicki J. (2008). Wpływ zmian klimatycznych na kształtowanie się zagrożenia pożarowego lasów w Polsce. Leśne Prace Badawcze, 69 (1).

15. Tomczyk K. (2003). Ochrona p.poż. – co się opłaca. Las Polski, 4.
16. Wilkowski W. Metoda określenia wartości lasów oraz gruntów zadrzewionych i zakorzenionych wchodzących w obszar scalania. (2002). Przegląd Geodezyjny, 8.
17. Wiler K., Wiśniewski W. Zabezpieczenie przeciwpożarowe lasów. (2008). Las Polski, 4.
18. Wiler K., Wiśniewski W. Zabezpieczenie przeciwpożarowe lasów. (2003). Las Polski, 4.

FIRE INSURANCE FOR FORESTS IN THE POLISH ZONE CLIMATE

Summary

Forests are a national wealth. They are exposed to fires caused by lightning, technical devices and human.

There is a lack of financial resources for forest management.

The article proposes introducing the forest insurance from the fires caused by nature. It points the principles of fire damage evaluation, frequency of fire occurrence and basic assumption for such insurances.

MISCELLANEA

Sławomir Biruk

Piotr Jaśkowski

Politechnika Lubelska

OCENA SKUTECZNOŚCI WYBRANYCH METOD REDUKCJI WARIANCJI W BADANIACH SYMULACYJNYCH MODELI SIECIOWYCH PRZEDSIĘWZIĘĆ BUDOWLANYCH*

Wprowadzenie

Analiza modeli sieciowych przedsięwzięć metodą symulacji cyfrowej Monte Carlo ma na celu najczęściej ustalenie średnich, wariancji lub typu rozkładu terminów zaistnienia zdarzeń i ($i = 1, 2, \dots, m$) na podstawie obserwacji $x_{i,j}$ poczynionych w j ($j = 1, 2, \dots, n$) przebiegach symulacyjnych. Wartości $x_{i,j}$ można traktować jako realizacje zmiennych losowych $X_{i,j}$ terminów zaistnienia zdarzeń [9].

Istota metody *MC* w analizie modeli sieciowych polega na losowym generowaniu w kolejnych przebiegach symulacyjnych czasów realizacji $t_{l,j}$ czynności ($l = 1, 2, \dots, w$, w – liczba czynności modelu sieciowego) i obliczaniu terminów zaistnienia zdarzeń jak w Critical Path Method (CPM). W każdej replikacji j czasy realizacji czynności są generowane zgodnie z przyjętym rozkładem prawdopodobieństwa, przy zastosowaniu niezależnych od siebie ciągów liczb losowych $u_{l,j}$ z przedziału $(0, 1]$.

Nieobciążony estymator punktowy średniej $\mu_i = E(X_i)$ zmiennej losowej X_i terminu zaistnienia zdarzenia i ma postać [5; 9]: $\bar{\mu}_i = \frac{1}{n} \sum_{j=1}^n X_{i,j}$, natomiast

* Praca została sfinansowana przez Ministerstwo Nauki i Szkolnictwa Wyższego (grant nr N N506 254637).

jego wariancja dana jest wzorem: $D^2(\bar{\mu}_i) = \frac{\sigma_i^2}{n}$, gdzie σ_i^2 jest wariancją zmiennej losowej X_i terminu zaistnienia zdarzenia i , tj. $\sigma_i^2 = D^2(X_i)$.

Nieobciążony estymator wariancji jest określony następująco

$$\bar{\sigma}_i^2 = \frac{1}{n-1} \sum_{j=1}^n (X_{i,j} - \bar{\mu}_i)^2$$

W metodzie symulacji *MC* błąd estymacji średniej $(\bar{\mu}_i - \mu_i)$ jest przedstawiany w postaci przedziału ufności. Wielkość przedziału ufności jest proporcjonalna do $\sigma_i / \sqrt{n}$. Jednym z podstawowych sposobów zawężania przedziału ufności (redukcji wariancji estymatora średniej) jest zwiększanie liczby obserwacji (przebiegów symulacyjnych), co jednak powoduje wydłużenie czasu badań symulacyjnych.

Na dokładność wyznaczanych charakterystyk wpływa także sposób generowania zmiennych (danych) wejściowych do badań. Wywołując korelację między wartościami tych samych wielkości w eksperymentach symulacyjnych, uzyskuje się redukcję wariancji estymatora, a co się z tym wiąże zwiększa się dokładność oszacowania. Zawężenie przedziału ufności estymowanej charakterystyki można także uzyskać dokonując właściwego wyboru estymatora.

Postępowania mające na celu zmniejszenie „rozrzutu” obserwowanych wartości zmiennych wyjściowych, w literaturze określa się mianem metod redukcji wariancji [6; 12]. W opracowaniu podjęto próbę oceny efektywności metod redukcji wariancji, polegających na zmianie sposobu generowania liczb losowych (metoda zmiennych antytetycznych, losowania warstwowego i metoda Quasi-Monte Carlo). Analizowane metody nie wydłużają czasu eksperymentów i są łatwe do zaimplementowania w komercyjnych językach symulacyjnych.

1. Metoda zmiennych antytetycznych

Metoda ta, zwana także losowaniem przeciwnym, polega na przeprowadzeniu eksperymentu symulacyjnego w dwóch etapach w sposób następujący [6; 12]. W pierwszym etapie stosuje się dla każdej czynności ciąg liczb pseudolosowych $u_{l,1}, u_{l,2}, \dots, u_{l,n/2}$ ($l = 1, 2, \dots, w$) o rozkładzie równomiernym na przedziale $(0, 1]$, natomiast w drugim – ciąg liczb dopełniających $(1 - u_{l,1}), (1 - u_{l,2}), \dots, (1 - u_{l,n/2})$. Jeżeli $\hat{X}_i^{(1)}$ oraz $\hat{X}_i^{(2)}$ są odpowiednio

estymatorami parametru $\hat{X}_i$ (średniego terminu zaistnienia zdarzenia i w modelu sieciowym) odpowiednio w pierwszym i drugim etapie badań symulacyjnych, to nieobciążonym estymatorem tego parametru jest średnia:

$\hat{X}_i = \frac{1}{2}(\hat{X}_i^{(1)} + \hat{X}_i^{(2)})$. Wariancję tego estymatora można obliczyć następująco

$$D^2(\hat{X}_i) = D^2\left(\frac{1}{2}(\hat{X}_i^{(1)} + \hat{X}_i^{(2)})\right) = \frac{1}{4}(D^2(\hat{X}_i^{(1)}) + D^2(\hat{X}_i^{(2)})) + \frac{1}{2}\text{cov}(\hat{X}_i^{(1)}, \hat{X}_i^{(2)})$$

Stosując losowanie przeciwstawne, z ujemną korelacją między wynikami z następujących po sobie eksperymentów, uzyskuje się mniejszą wariancję niż w przypadku prowadzenia eksperymentów niezależnych.

Burt i inni [3] zastosowali ciągi liczb antytetycznych do symulacyjnego rozwiązywania prostych modeli sieciowych z czasami realizacji procesów opisanymi za pomocą rozkładu wykładniczego. Następnie wykazali analitycznie, że wariancję oczekiwanego czasu realizacji modelowanych przedsięwzięć można zmniejszyć o ponad połowę w stosunku do wariancji estymatora stosowanego w bezpośredniej metodzie Monte Carlo. Rozważania te zostały uogólnione dla różnych typów rozkładów teoretycznych czasów realizacji czynności w pracy Sullivana i innych [11]. Stosując losowanie przeciwstawne, tę samą dokładność co w metodzie MC można uzyskać średnio przy liczbie replikacji mniejszej o 75%. W przypadku, gdy rozkłady czasów są symetryczne, można zakładać, że – w celu uzyskania tej samej dokładności – liczba powtórzeń może zostać zmniejszona o 90%.

2. Metoda Quasi-Monte Carlo (QMC)

Cechą charakterystyczną liczb pseudolosowych, generowanych w symulatorach, jest tendencja do ich grupowania. W celu uniknięcia tego zjawiska, podczas badań symulacyjnych można stosować ciągi (sekwencje) liczb quasi losowych $\{x_{i,j}\}$. Liczby te są rozłożone w sposób bardziej równomierny niż pseudolosowe [4].

Metoda QMC różni się od metody podstawowej jedynie sposobem generowania liczb, co czyni ją łatwą do implementacji w językach symulacyjnych. Wyniki zastosowań tej metody w analizie sieci PERT są prezentowane w pracy [13].

3. Metoda losowania warstwowego

W metodzie losowania warstwowego dystrybucje rozkładów zmiennych losowych czasu wykonania wszystkich czynności l ($l = 1, 2, \dots, w$) są dzielone na $s(l)$ rozłącznych przedziałów (warstw) $[a_{l,k}, a_{l,k+1}]$ takich, że $0 = a_{l,1} < a_{l,2} < \dots < a_{l,s(l)-1} = 1$. Liczba przedziałów nie musi być jednakowa dla każdej czynności, ale ich równa liczba ułatwia planowanie i prowadzenie badań symulacyjnych. Najlepsze rezultaty uzyskuje się dokonując takiego podziału, aby wariancje w każdej warstwie były sobie równe. Najczęściej jednak – ze względu na łatwość implementacji – dokonuje się podziału wszystkich dystrybucji zmiennych losowych wejściowych na jednakową liczbę s przedziałów o jednakowych długościach.

W ogólnym przypadku, aby dokładnie odzorować zmienną losową będącą wynikiem symulacji, należy przeprowadzić $s(1) \cdot s(2) \cdot \dots \cdot s(w)$ replikacji, losując z każdej warstwy zmiennej wejściowej po jednej wartości. Takie podejście jest efektywne tylko przy stosunkowo niewielkiej liczbie czynności modelu sieciowego.

Przy dużej liczbie czynności w modelu eksperyment symulacyjny można przeprowadzić stosując schemat Latin Hypercube Sampling (LHS), zaproponowany w pracy [7]. Aby wygenerować s ($j = 1, 2, \dots, s$) skorelowanych ze sobą replikacji (liczba replikacji musi być równa liczbie warstw), wartości liczb losowych $U_{l,j}$ (do ustalania realizacji zmiennych losowych) należy określać następująco [8]

$$U_{l,j} = \frac{\pi_l(j) - 1 + U_{l,j}^*}{s} \quad \text{dla } l = 1, 2, \dots, w \text{ oraz } j = 1, 2, \dots, s$$

gdzie:

$\pi_1(\cdot), \pi_2(\cdot), \dots, \pi_w(\cdot)$ – są permutacjami liczb $\{1, 2, \dots, s\}$, losowanymi z jednakowym prawdopodobieństwem ze zbioru $s!$ takich permutacji, a $\pi_l(j)$ oznacza j -ty element (numer warstwy) w permutacji dla czynności l ,

$\{U_{l,j}^* : l = 1, 2, \dots, w \text{ oraz } j = 1, 2, \dots, s\}$ – są liczbami losowymi wzajemnie niezależnymi oraz niezależnymi od permutacji $\pi_1(\cdot), \pi_2(\cdot), \dots, \pi_w(\cdot)$.

W każdym cyklu r ($r = 1, 2, \dots, k$) badań symulacyjnych, obejmującym s replikacji ustalonych według procedury LHS, jest losowana tylko jedna liczba z każdej warstwy dla danej czynności. Nieobciążonymi estymatorami wartości średnich $\mu_i = E(X_i)$ zmiennych losowych X_i terminów zaistnienia zdarzeń

w modelu sieciowym są w każdym cyklu $\overline{X_{i,r}} = \frac{1}{s} \sum_{j=1}^s X_{i,r,j}$ [1].

Przeprowadzenie n przebiegów symulacyjnych wymaga wygenerowania $k = \frac{n}{s}$ niezależnych permutacji $\pi_1(\cdot), \pi_2(\cdot), \dots, \pi_w(\cdot)$. W każdym powtórzeniu są obliczane w metodzie *CPM* wartości zmiennych losowych zaistnienia zdarzeń w modelu sieciowym, a więc w konsekwencji estymatorami wartości średnich tych zmiennych są także: $\overline{X_i} = \frac{1}{k} \sum_{r=1}^k X_{ir}$ [2]. Stosując metodę *LHS* zawsze uzyskuje się redukcję wariancji estymatora w stosunku do wariancji uzyskanej w klasycznej metodzie *MC*.

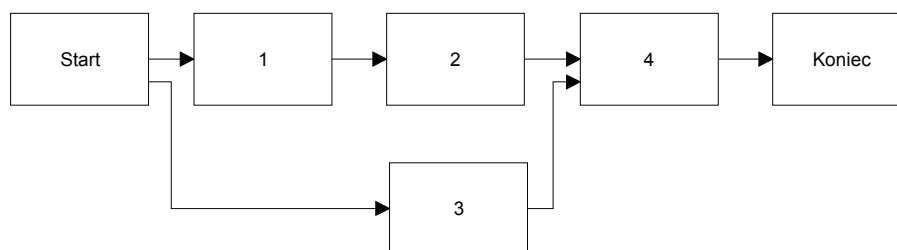
W literaturze dotyczącej symulacji systemów dyskretnych [12] jest prezentowana także szczególna odmiana metody losowania warstwowego, podobna do losowania przeciwnastawnego. Eksperyment symulacyjny jest powtarzany dwukrotnie. W pierwszym etapie są stosowane niezależne dla czynności ciągi liczb pseudolosowych $u_{l,1}, u_{l,2}, \dots, u_{l,n/2}$ o rozkładzie równomiernym na przedziale $(0, 1]$, w drugim etapie – ciągi spełniające warunek

$$u_{l,j}^* = \begin{cases} u_{l,j} + 0,5 & \text{dla } 0 \leq u_{l,i} < 0,5 \\ u_{l,j} - 0,5 & \text{dla } 0,5 \leq u_{l,i} \leq 1 \end{cases}$$

Pomiędzy zmiennymi losowymi ustalonymi na podstawie ciągów liczb $u_{l,j}$ oraz $u_{l,j}^*$, $j = 1, 2, \dots, n/2$ wywołuje się ujemną korelację i w efekcie zmniejsza się wariancję estymatora wartości średniej.

4. Badania symulacyjne z zastosowaniem różnych metod

Przeprowadzono badania symulacyjne dwóch modeli sieciowych (rys. 1 i 2), w których czasy wykonania czynności opisano za pomocą rozkładów trójkątnych o wartościach minimalnych a_l , maksymalnych b_l oraz najbardziej prawdopodobnych c_l . Parametry rozkładów zestawiono w tabelach 1 i 2. Wszystkie programy symulacyjne zostały opracowane w języku symulacyjnym GPSS World (Minuteman Software). Dla każdego wariantu badań symulacyjnych przeprowadzono łącznie 1200 replikacji (obliczeń czasu trwania przedsięwzięcia dla 1200 zestawów wartości zmiennych losowych czasów trwania czynności modelu), co zapewniło porównywalność efektywności zaimplementowanych metod.



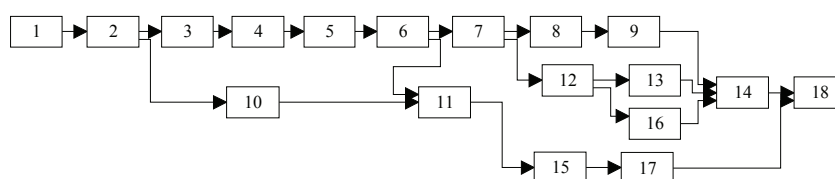
Rys. 1. Graf zależności kolejnościowych (przykład 1)

Tabela 1

Parametry zmiennych losowych czasu wykonania czynności modelu sieciowego
(przykład 1) [dni]

Numer czynności	a_l	c_l	b_l
Start	0	0	0
1	5	6	8
2	4	5	7
3	8	10	15
4	3	4	7
Koniec	0	0	0

Źródło: Opracowano na podstawie własnych oszacowań.



Rys. 2. Graf zależności kolejnościowych (przykład 2)

Tabela 2

Parametry zmiennych losowych czasu wykonania czynności sieci (przykład 2) [dni]

l	Czynność (proces budowlany)	a_l	c_l	b_l
1	Rozpoczęcie	0	0	0
2	Zagospodarowanie placu budowy	5	6	8
3	Wykopy	4	5	7
4	Fundamenty	8	10	15
5	Ściany piwnic	3	4	7
6	Podłoża i posadzki w piwnicy	3	4	7
7	Konstrukcja budynku	15	18	25
8	Konstrukcja dachu	7	8	10
9	Pokrycie dachu	5	6	8
10	Przyłącza	2	3	4
11	Obsypanie fundamentów	1	2	4
12	Ściany działowe	3	5	8
13	Tynki	8	10	15
14	Malowanie	5	6	9
15	Roboty elewacyjne	10	12	18
16	Instalacje wewnętrzne	10	12	15
17	Zagospodarowanie terenu	14	16	22
18	Zakończenie	0	0	0

Źródło: Opracowano na podstawie własnych oszacowań.

Wyniki badań modelu sieciowego z rys. 1 zestawiono w tabeli 3, natomiast z rys. 2 – w tabeli 4.

W pierwszym etapie badań symulacyjnych stosowano predefiniowane w języku GPSS World generatory rozkładu trójkątnego w postaci $TRIANGULAR(RN, a, c, b)$, gdzie RN jest numerem generatora liczb losowych. Przypisanie kolejnym czynnościom generatorów o różnych numerach zapewnia niezależność zmiennych losowych czasu wykonania poszczególnych czynności.

W metodzie Quasi-Monte Carlo zastosowano sekwencję Weyla – liczb pseudolosowych definiowanych w następujący sposób

$$x_v = v \cdot \Theta - \lfloor v \cdot \Theta \rfloor; v = 1, 2, \dots, n$$

gdzie Θ – jest podstawą ciągu Weyl'a (dowolną liczbą niewymierną z przedziału $[0, 1)$).

Do generowania czasów realizacji kolejnych czynności stosowano różne podstawy ciągu Weyl'a Θ , np. $\Theta_1 = \frac{\sqrt{3}-1}{2}$, $\Theta_2 = \frac{\sqrt{5}-1}{2}$, ...

Predefiniowane w języku symulacyjnym GPSS World generatory rozkładów, w tym rozkładu trójkątnego, nie pozwalają na bezpośrednie stosowanie niektórych metod redukcji wariancji, np. losowania przeciwnego. Dlatego w kolejnych wariantach symulatora do generowania wartości zmiennych losowych o rozkładzie trójkątnym z parametrami (a_j, b_j, c_j) zastosowano metodę odwracania dystrybucji:

$$F^{-1}(u | a_j, b_j, c_j) = \begin{cases} a_j + \sqrt{u(c_j - a_j) \cdot (b_j - a_j)} & \text{dla } 0 \leq u \leq \frac{c_j - a_j}{b_j - a_j} \\ b_j - \sqrt{(1-u) \cdot (b_j - c_j) \cdot (b_j - a_j)} & \text{dla } \frac{c_j - a_j}{b_j - a_j} \leq u \end{cases}$$

gdzie u – jest liczbą losową o rozkładzie równomiernym na przedziale $(0, 1]$.

Procedura generująca wartości zmiennej losowej o rozkładzie trójkątnym (czasy wykonania czynności) w języku GPSS ma postać

PROCEDURE ROZKLAD() BEGIN

LOS=LOS/1000

IF (LOS<=(C-A)/(B-A)) THEN BEGIN

CZAS = (A+ SQR(LOS*(C-A)*(B-A)));

RETURN CZAS;

END;

ELSE BEGIN

CZAS = (B- SQR((1-LOS)*(B-C)*(B-A)));

RETURN CZAS;

END;

END;

W przypadku zastosowania metody losowania przeciwnego w powyższej procedurze należy dokonać podstawienia: $LOS:=1-LOS$. Różne numery generatorów LOS liczb losowych (z zakresu 0-1000) zapewniają niezależność zmiennych losowych.

W metodzie losowania Latin Hypercube Sampling zakres zmienności wszystkich zmiennych losowych został podzielony na pięć warstw o jednakowym prawdopodobieństwie wystąpienia. Permutacje losowe liczb naturalnych 1, 2, ..., 5 (warstw), każda o jednakowym prawdopodobieństwie wystąpienia, były generowane zgodnie z algorytmem prezentowanym w pracy [10].

W drugiej, szczególnej odmianie metody losowania warstwowego procedura generująca czasy wykonania czynności miała następującą postać

```
PROCEDURE ROZKLAD() BEGIN
```

```
    LOS=LOS/1000
    IF (LOS < 0.5) THEN BEGIN
        LOS = (LOS + 0.5);
    END;
    ELSE BEGIN
        LOS = (LOS - 0.5);
    END;
    IF (LOS<=(C-A)/(B-A)) THEN BEGIN
        CZAS = (A+ SQR(LOS#(C-A)#(B-A)));
        RETURN CZAS;
    END;
    ELSE BEGIN
        CZAS = (B- SQR((1-LOS)#(B-C)#(B-A)));
        RETURN CZAS;
    END;
END;
```

Podsumowanie

Metoda symulacji Monte Carlo jest efektywnym narzędziem analizy sieci zależności czynności o dowolnych rozkładach czasu ich trwania, bez konieczności wprowadzania dodatkowych założeń upraszczających. W ten sam sposób mogą być analizowane modele proste (obejmujące jedynie zależności kolejnościowe), jak i modele z ograniczeniami w dostępności zasobów lub ograniczeniami czasowymi (np. polityka railway). O jakości uzyskanych rezultatów decyduje sposób prowadzenia badań symulacyjnych. Zastosowanie metod redukcji wariancji może zarówno skrócić czas prowadzenia badań (zmniejszenie liczby replikacji), jak i podnieść wiarygodność oszacowania wybranych charakterystyk badanego modelu.

Tabela 3

Wyniki uzyskane przy stosowaniu różnych metod redukcji wariancji w badaniach modelu sieciowego z rys. 1

Metoda	Średnia	Wariancja estymatora
GPSS	16,758	0,001388
Metoda odwracania dystrybucyj	16,701	0,001450
Metoda Quasi-Monte Carlo	16,763	0,001361
Metoda losowania przeciwnego	16,752	0,000401
Losowanie warstwowe – metoda <i>LHS</i>	16,748	0,000239
Losowanie warstwowe – metoda 2	16,729	0,000872

Źródło: Opracowano na podstawie wyników badań symulacyjnych.

Tabela 4

Wyniki uzyskane przy stosowaniu różnych metod redukcji wariancji w badaniach modelu sieciowego z rys. 2

Metoda	Średnia	Wariancja estymatora
GPSS	76,007	0,010228
Metoda odwracania dystrybucyj	76,009	0,010588
Metoda Quasi-Monte Carlo	75,952	0,009466
Metoda losowania przeciwnego	75,951	0,000552
Losowanie warstwowe – metoda <i>LHS</i>	75,964	0,001069
Losowanie warstwowe – metoda 2	76,003	0,006340

Źródło: Opracowano na podstawie wyników badań symulacyjnych.

Efektywność poszczególnych metod redukcji wariancji zależy przede wszystkim od konfiguracji sieci zależności i stosowanych typów oraz parametrów rozkładów czasu trwania czynności, ale także od doświadczenia i umiejętności badacza.

Na podstawie przeprowadzonych eksperymentów symulacyjnych można wnioskować, że zastosowanie metody liczb antytetycznych prowadzi do znacznej redukcji wariancji. Porównywalną efektywnością charakteryzuje się metoda losowania warstwowego Latin Hypercube Sampling, ale w tym przypadku zaprojektowanie symulatora i przeprowadzenie badań symulacyjnych jest bardziej złożone. Kierunkiem przyszłych badań będzie analiza większej liczby modeli testowych oraz stosowanie łączne kilku metod redukcji wariancji, np. metody *LHS* i losowania przeciwnego.

Literatura

1. Avramidis A.N., Wilson J.R. (1996). Integrated Variance Reduction Strategies for Simulation. *Operational Research*, 44(2), s. 327-346.
2. Blower S.M., Dowlatabadi H. (1994). Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: an HIV Model, as an Example. *International Statistical Review*, 63(2), s. 229-243.
3. Burt J.M., Gaver D.P., Perlas M. (1970). Simple Stochastic Networks: Some Problems and Procedures. *Naval Research Logistics Quarterly*, 17(4), s. 439-459.
4. Gentle J.E. (2003). *Random Number Generation and Monte Carlo Methods*. Springer, New York.
5. Greń J. (1984). *Statystyka matematyczna. Modele i zadania*. Państwowe Wydawnictwo Naukowe, Warszawa.
6. Law A.M., Kelton W.D. (1991). *Simulation Modeling & Analysis*. McGraw Hill, International Edition.
7. McKay M.D., Beckman R.J., Conover W.J. (1998). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(2), s. 239-245.
8. Owen A.B. (1998). Latin Supercube Sampling for Very High-Dimensional Simulations. *ACM Transactions on Modeling and Computer Simulation*, 8(1), s. 71-102.
9. Platt C. (1974). *Problemy rachunku prawdopodobieństwa i statystyki matematycznej*. PWN, Warszawa.
10. Skiena S.S. (2008). *The Algorithm Design Manual*. Springer-Verlag, London.
11. Sullivan R.S., Hayya J.C., Schaul R. (1982). Efficiency of the Antithetic Variate Method for Simulating Stochastic Networks. *Management Science*, 28(5), s. 573-572.
12. Tyszer J. (1990). *Symulacja cyfrowa*. Wydawnictwa Naukowo-Techniczne, Warszawa.
13. Yang H. (2006). Quasi Random Sampling for Operations Management. *Seoul Journal of Business*, 12(1), s. 53-72.

ASSESSING EFFICIENCY OF VARIANCE REDUCTION METHODS IN CONSTRUCTION PROJECTS' NETWORK SIMULATION

Summary

Monte Carlo simulation is a popular tool that supports planning projects affected by risk. Analysing the results of computer simulations enables the planner to formulate and verify hypotheses on distribution type and parameters of schedule events occurrence and the project's duration. Simulation research on network models, whose structure is deterministic, but tasks times

are of stochastic nature, does not require introducing any simplifying assumptions on task duration distribution types. It also allows the planner to account for resource constraints, in terms of both number of available units and time of their availability.

Accuracy of estimates obtained by means of simulations can be improved by increasing the number of replications, or by applying variance reduction methods. The latter may consist in change of the way the random numbers are generated.

The paper analyses how the method of variance reduction affects simulation results in terms of standard error of estimated project duration mean value. The methods considered were: Quasi-Monte Carlo, stratified sampling, antithetic variates. The object analysis were network models with task durations of triangular distribution. This type of distribution is commonly assumed in modelling the effect of random occurrences on organisation of construction works.

Urszula Grzybowska

Marek Karwański

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

PRZYKŁADY I PORÓWNANIE MODELI MACIERZY MIGRACJI STOSOWANYCH W ANALIZIE RYZYKA KREDYTOWEGO

Wprowadzenie

Od końca lat 80. środowiska bankowe dążyły do uporządkowania nie tylko procedur kredytowych, ale i metod kwantyfikacji ryzyka finansowego z nimi związanego. Proces zarządzania ryzykiem kredytowym powiązano z podstawowym celem, jakim była realizacja zrównoważonego wzrostu, zgodnego z założonymi wynikami finansowymi. Niezbędne stały się modele pozwalające na prawidłową wycenę ryzyka. Podstawą modeli są koszty generowane przez ryzyko, odzwierciedlane np. w spreadach wycen rynkowych. Początkowo dominowało podejście strukturalne – modelowano procesy zmian wartości aktywów i pasywów badanych kredytobiorców w ramach teorii Merton. Pozwalało to zdefiniować wielkość zwaną „odległością od defaultu” i powiązać ją z wartością kredytobiorcy. W latach 90. większą uwagę zwrócono na rolę systemów ratingowych mierzących „zdolność kredytową” i procesy zmian ratingów w czasie. Opis tych zmian wyrażano w postaci intensywności Λ

$$P(\text{default} \in [t, t + \Delta t] \mid H_t) \approx \lambda(t)\Delta t$$

gdzie H_t oznacza informację historyczną.

Za stosowaniem tych modeli przemawiają obserwacje wycen rynkowych klientów mających jednakowe ratingi [12]. Okazuje się, że można powiązać intensywność Λ bezpośrednio ze spreadami krzywych dochodowości obliczanych dla tych klas klientów, a co za tym idzie z ich rynkową wyceną.

Najprostszą metodą opisaną intensywności jest wprowadzenie elementów skończonej przestrzeni stanów łańcuchów Markowa [5; 9]. Procedury kwantyfikacji intensywności oparte zostały na obserwowanych macierzach przejścia (migracji) kredytobiorcy. Istnieje wiele metod estymacji macierzy migracji. Najczęściej używa się metod częstościowych opartych na bezpośrednich obserwacjach danych historycznych. Inną możliwość otwierają teorie statystyczne w ramach longitudinalnych modeli liniowych GLMM.

1. Cel i metoda badawcza

W niniejszej pracy porównujemy wpływ stosowania różnych estymatorów macierzy migracji na pomiar ryzyka w banku. Obecnie w praktyce stosuje się cztery modele: Model Mertona, CreditRisk+, CreditMetrics, Credit-Portfolio View. Ostatnie dwa z powyższych modeli oparte są na macierzach migracji uzyskiwanych na podstawie obserwacji częstości zdarzeń historycznych.

W krajach o rozwiniętej infrastrukturze odpowiednie estymatory macierzy migracji publikowane są przez zewnętrzne agencje ratingowe, np. S&P [20], Moody's [15; 16]. Dane te są wykorzystywane w bankach do modelowania prawdopodobieństw defaultów. Główną przeszkodą stosowania tych metod w Polsce jest brak odpowiednich danych. Liczba przedsiębiorstw ratowanych przez zewnętrzne agencje w naszym kraju jest niewielka. Banki muszą opierać się na swoich wewnętrznych modelach. Niestety, pojedyncze banki dysponują małą liczbą danych o defaultach. Dodatkowo dane historyczne są często niskiej jakości. Wiąże się to np. z brakiem spójnej definicji defaultu: klienci, którzy mają kłopoty ze spełnieniem warunków finansowych podejmują próby restrukturyzacji kredytów. Tymczasem bardzo często nie jest to uwidocznione w danych. W takich sytuacjach moment zajścia zdarzenia „default” nie może być precyzyjnie określony. Kolejnym elementem utrudniającym korzystanie z baz danych banku jest fakt, że wewnętrzne systemy ratingowe budowane są zazwyczaj opierając się na metodyce PIT (Point-in-Time), czyli dla konkretnego stanu gospodarki.

Z uwagi na powyższe uwarunkowania przyjęliśmy za celowe porównanie metod na podstawie danych generowanych sztucznie. Takie podejście obciążone jest innymi obciążeniami, z których najważniejsze jest to, że generator oparty jest na stacjonarnym rozkładzie wielomianowym, który rzadko występuje w tej postaci w praktyce. W rezultacie uzyskane wyniki nie powinny służyć do oceny modeli pod kątem dokładności dopasowania (goodness-of-fit), ale do stwierdzenia czy różnice pomiędzy modelami są rzeczywiście duże i do zdefiniowania warunków stosowalności tych modeli.

2. Wyznaczanie macierzy migracji na podstawie modeli łańcuchów Markowa

W pracy będziemy opierać się na ratingu zaproponowanym przez [15], obejmującym 8 stanów. Opis sposobu klasyfikacji kredytobiorcy do jednej z klas ratingowych oraz specyfikację klas można znaleźć w dyrektywach ba-
zylejskich [2] lub w [18].

W spotykanych w literaturze tematu rozważaniach oraz w zastosowaniach modeli przyjmuje się, że zmiany ratingów można opisać za pomocą skończonego, jednorodnego łańcucha Markowa o macierzy przejścia $P = [p_{ij}]$ wyznaczonej na podstawie danych empirycznych jednym z kilku sposobów. Niestety, założenie jednorodności lub inaczej niezmienniczości w czasie nie jest zazwyczaj spełnione. Brak niezmienniczości w czasie zauważony już w [6; 7], spowodowany jest najczęściej czynnikami zewnętrznymi (np. okresy cyklu gospodarczego). Zaproponowany w [8] test niezmienniczości w czasie wymaga dostatecznie dużej próby dla każdego wyrazu macierzy, a jak wiadomo, migracje poza przekątną macierzy są stosunkowo nieliczne. Z tego też powodu elementy leżące na diagonalu są szacowane z dużą dokładnością, zaś elementy poza przekątną, z dużym błędem. Szczególną uwagę należy zwrócić na zera występujące w macierzach migracji. Zero oznacza brak albo niemożliwość migracji, a to z punktu widzenia zastosowań nie jest pożądane. Z punktu widzenia zastosowań w ocenie ryzyka kredytowego bardzo ważna jest ostatnia kolumna macierzy migracji, która przedstawia prawdopodobieństwa migracji do stanu niewypłacalności (default). Macierze przejścia stosowane do opisu migracji klientów są diagonalnie dominujące, nieujemne, stochastyczne (suma elementów w wierszu wynosi 1), mają największą wartość własną równą 1, ich pozostałe wartości własne są bliskie 1, posiadają stan pochłaniający (default).

Znanych jest kilka metod estymacji macierzy przejścia dla danych indywidualnych. Najczęściej stosowanymi metodami są: metoda kohort, metoda okresowa parametryczna (tzw. metoda duration) oraz metoda okresowa nieparametryczna (Nelsona-Aalena-Johansena), która nie wymaga założenia niezmienniczości w czasie. W pracy pokażemy, że macierze przejścia można także szacować wykorzystując longitudinalne uogólnione modele liniowe (GLMM).

2.1. Metoda Kohortowa

Niech n_{ij} oznacza liczbę elementów (kredytobiorców), które były w stanie i w okresie $t-1$ oraz w stanie j w okresie t . Możemy oszacować prawdopodobieństwo, że kredytobiorca znajdzie się w stanie j w okresie t pod warunkiem, że był w stanie i w okresie $t-1$, wykorzystując wzór

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}} \quad (1)$$

Stąd prawdopodobieństwo przejścia z dowolnego stanu i do stanu j jest równe stosunkowi liczby elementów n_{ij} , które przeszły ze stanu i do stanu j , do ilości wszystkich elementów $\sum_j n_{ij} = N_i$, które wyszły ze stanu i .

Metoda kohortowa [11; 13; 14] dla danych indywidualnych wymaga śledzenia losu każdego „klienta”. W algorytmie prawdopodobieństwa przejścia obliczane są na podstawie ratingów dla dwóch wybranych dat. Nie są brane pod uwagę zmiany ratingu między okresami. Metoda kohortowa jest metodą najczęściej stosowaną w praktyce.

2.2. Estymacja macierzy przejścia z uwzględnieniem T okresów

W opracowaniu [19] zaproponowano następujący estymator prawdopodobieństw przejścia z uwzględnieniem T okresów

$$p_{ij} = \frac{N_{ij}}{N_i} = \frac{\sum_{t=1}^T n_{ij}(t)}{\sum_{t=1}^T N_i(t)}, \text{ gdzie } N_i = \sum_j n_{ij} \quad (2)$$

zaś $N_i(t)$ jest wartością sumy dla okresu t .

Estymator ten w dalszej części pracy nazwiemy Plug-in, a macierz przejścia wyznaczoną według wzoru (2) oznaczmy przez PI.

2.3. Estymator Duration

W przeciwieństwie do metody kohortowej, algorytm estymacji „duration” wykorzystuje pełną historię ratingową [14; 17].

Przy założeniu niezmienniczości w czasie, prawdopodobieństwa przejścia mogą być opisane przy pomocy macierzy intensywności Λ . Macierz przejścia $P(t)$, dla dowolnego $t \geq 0$, może być wyznaczona według wzoru

$$P(t) = \exp(\Lambda(t)) = I + t\Lambda(t) + t^2 \frac{\Lambda^2(t)}{2!} + t^3 \frac{\Lambda^3(t)}{3!} + \dots \quad (3)$$

$t \geq 0$, gdzie wyrazy macierzy Λ spełniają warunek $\lambda_{ij} \geq 0$, dla $i \neq j$ oraz $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$. Estymator macierzy Λ jest dany wzorem

$$\lambda_{ij} = \frac{n_{ij}(T)}{\int_0^T Y_i(s) ds}$$

gdzie $Y_i(s)$ jest liczbą kredytobiorców w stanie i , w chwili s , zaś $n_{ij}(T)$ jest sumą przejść ze stanu i do j dla $i \neq j$ w czasie do T .

Podejście, które nie wymaga założenia niezmienniczości w czasie oparte jest na teorii Aalena-Johansena znanej także jako metoda Nelsona-Aalena. Estymator macierzy przejścia w okresie czasu od s do t , jest iloczynem macierzy danym wzorem

$$P(s, t) = \prod_{i=1}^m (I + \Delta \hat{A}(T_i))$$

gdzie wyrazy poza przekątną macierzy $\Delta \hat{A}(T_i)$ są postaci

$$a_{jk} = \frac{\Delta N_{jk}(T_i)}{Y_j(T_i)} \quad (4)$$

przy czym T_i oznacza skok (przedział czasowy) w okresie od s do t . $Y_j(T_i)$ jest liczbą kredytobiorców znajdujących się w stanie j na początku okresu, licznik: $\Delta N_{jk}(T_i)$ oznacza ilość przejść ze stanu j do k w okresie T_i , m jest liczbą dni, w których zdarzyła się przynajmniej jedna zmiana stanu [14; 17]. Suma wyrazów we wierszach równa się 0.

3. Zastosowanie Uogólnionych Mieszanych Modeli Liniowych (GLMM) do wyznaczanie macierzy migracji

Ponieważ prawdopodobieństwa przejścia $p_{ij} = P(Y_t = j | Y_{t-1} = i)$ można rozpatrywać z punktu widzenia statystyki jako zdarzenia powiązane (matched pairs), wydaje się, że naturalne byłoby zastosowanie modeli longitudinalnych GLMM. Modele te stanowią rozszerzenie uogólnionych modeli liniowych GLM [1]. W niniejszym opracowaniu wybraliśmy dwa typy modeli GLMM: model należący do klasy modeli brzegowych (Marginal Models) oraz model przejść (Transitional Models).

3.1. Modele Brzegowe (Marginal GLMM)

Modele brzegowe to klasa modeli należących do uogólnionych modeli liniowych, w których wartość średnią i wariancję definiuje się niezależnie. Niech funkcja gęstości opisująca prawdopodobieństwo $f(\cdot)$ będzie należała do jednoparametrycznej rodziny wykładniczej [1, 3]. W modelach brzegowych wariancja $Var(Y_{ij})$ zależy od wartości średniej brzegowej poprzez zadaną funkcję. Tak zdefiniowana klasa modeli nosi nazwę modeli brzegowych [1]. W niniejszej pracy do modelowania macierzy migracji (8 stanów w tym stan pochłaniający) wybrano model brzegowy, składający się z wartości średniej i wariancji równych sobie. Wartość średnia opisana jest równaniem, w którym zmienne egzogeniczne zależą od klas ratingowych

$$\log\{E(Y_{ij})\} = \alpha_i \cdot rating_i^t + \beta_j \cdot rating_j^{t-1} + \theta_{ij} \cdot factor_{ij} + \tau_{ij} \cdot \delta_{ij}$$

gdzie δ_{ij} oraz $factor_{ij}$ zależą od schematu ratingowego.

3.2. Modele Przejść (Transitional GLMM)

W modelach przejścia przyjmuje się założenia uogólnionych modeli liniowych z tym, że funkcja gęstości rozkładu $f(Y_{ij})$ jest warunkowa i zależy explicite od historii oraz dodatkowych zewnętrznych (egzogenicznych) zmiennych X_{ij} [1; 3]. Oznaczmy historię przez $H_{ij} = \{Y_{i1}, Y_{i2}, \dots, Y_{i,j-1}\}$. Wówczas rozkład $f(Y_{ij}|H_{ij})$ należy do klasy jednoparametrycznej rodziny wykładniczej. Warunkową średnią możemy przedstawić w postaci

$$g(E(Y_{ij}|H_{ij})) = X_{ij}\beta + \sum_{r=1}^s \alpha_r f_r(H_{ij})$$

gdzie funkcja łącząca $g(\cdot)$ oraz funkcje transformujące $f(\cdot)$ mają zadaną postać.

Na potrzeby niniejszej pracy do modelowania macierzy migracji (8 stanów w tym stan pochłaniający) wybraliśmy następujące równania, w których funkcje transformujące są identycznościowe, a zmienne egzogeniczne zależą od klas ratingowych

$$\log\{E(Y_{ij}|y_{ij;t-1})\} = \alpha_i \cdot rating_i^t + \beta_j \cdot rating_j^{t-1} + \gamma_{ij} y_{ij;t-1}$$

4. Generatory łańcuchów Markowa i ich znaczenie

Niezależnie od stosowanej metody estymacji, najkrótszym okresem, dla którego wyznacza się macierze przejścia jest jeden rok (okres). Dla krótszych przedziałów czasowych ilość migracji między stanami może być zbyt mała, aby uzyskane wyniki były wiarygodne. Mimo to w praktyce istnieje potrzeba wykorzystywania macierzy przejścia dla krótszych okresów. Narzędziem pozwalającym na wykorzystywanie macierzy przejścia $P(t)$ dla dowolnego $t \geq 0$, jest wyznaczenie generatora Q macierzy przejścia P takiego, że $P = \exp(Q)$. Wówczas podstawienie

$$P(t) = \exp(tQ) = I + tQ + \frac{(tQ)^2}{2!} + \frac{(tQ)^3}{3!} + \dots \quad (5)$$

daje macierz przejścia w dowolnym momencie $t \geq 0$. Innymi słowy wyznaczenie generatora Q macierzy przejścia jedno okresowej, a więc wyznaczonej dla czasu dyskretnego, pozwala na uzyskanie macierzy przejścia z czasem ciągłym [10].

Twierdzenie 1 [10]

Niech P będzie macierzą przejścia stopnia N .

Niech

$S = \max \{(a-1)^2 + b^2\}$; gdzie $a+ib$ ($a, b \in R$) jest wartością własną macierzy P . Załóżmy, że $S < 1$. Wówczas szereg

$$\tilde{Q} = (P-I) - \frac{(P-I)^2}{2} + \frac{(P-I)^3}{3} - \frac{(P-I)^4}{4} + \dots \quad (6)$$

jest zbieżny do macierzy $\tilde{Q}$ takiej, że $\exp(\tilde{Q}) = P$.

W rozważanych w pracy przykładach wartości własne są rzeczywiste, zaś największa wartość własna jest równa 1. Stąd generatory można obliczyć wykorzystując powyższy wzór. Warto zauważyć, że w [10] i [17] podano wiele twierdzeń i kryteriów istnienia i jednoznaczności generatorów macierzy przejścia.

Twierdzenie 2 [10]

Założmy, że wyrazy na przekątnej macierzy P są większe od $1/2$, tj. $p_{ii} > 0,5$ dla każdego i . Wówczas $S < 1$, czyli szereg (6) jest zbieżny.

Ponieważ leżące poza przekątną wyrazy generatora $\tilde{Q}$ mogą być ujemne, wymaga to wprowadzenia korekty, np. poprawki zdefiniowanej w [10]. Wyrazy generatora Q obliczamy jako modyfikację $\tilde{Q}$ według wzoru

$$q_{ij} = \begin{cases} \max(\tilde{q}_{ij}, 0), & \text{dla } j \neq i \\ \tilde{q}_{ii} + \sum_{j \neq i} \min(\tilde{q}_{ij}, 0), & i = j \end{cases} \quad (7)$$

Modyfikacja oznacza, że ujemne wartości poza przekątną dodajemy do wyrazu leżącego na przekątnej i zastępujemy je zerami. Zastosowana poprawka pozwala na zachowanie podstawowych własności generatora.

Po wprowadzeniu poprawek według wzoru (7) wyznaczana jest według wzoru (5) macierz przejścia P , dla dowolnego $t \geq 0$.

5. Opis danych

Do analizy porównawczej macierzy migracji wykorzystane zostały dane wygenerowane na podstawie rzeczywistej macierzy migracji [15]*.

Tabela 1

Macierz migracji

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,8933	0,1018	0,0036	0	0,0012	0	0	0
Aa	0,0086	0,8781	0,1065	0,0029	0,0019	0	0	0,0019
A	0	0,0153	0,9027	0,0658	0,0126	0,0027	0	0,0009
Baa	0	0,0052	0,0628	0,8356	0,0817	0,0126	0	0,0021
Ba	0,001	0	0,0029	0,0391	0,8418	0,1035	0,0039	0,0078
B	0	0,001	0,0019	0,0038	0,0596	0,8221	0,0298	0,0817
Caa_C	0	0	0	0,0138	0,0276	0,0741	0,6034	0,281
D	0	0	0	0	0	0	0	1

Źródło: [15].

Na jej podstawie wygenerowano 2000 hipotetycznych klientów z jednokowym prawdopodobieństwem w każdej klasie ratingowej („Aaa”, „Aa”, „A”, „Baa”, „Ba”, „B”, „Caa_C” oraz „D”), co miało odpowiadać schematowi badania typu case-control. Klasy ratingowe różnią się prawdopodobieństwem przejść. Następnie przeliczone zostały trajektorie 5 okresowe dla każdego klienta zgodnie z zadaną macierzą migracji. Dane te stanowiły bazę, na której szacowane były modele migracji.

* Zastosowanie opisanych metod na danych rzeczywistych wymaga spełnienia kilku warunków pozwalających na uznanie jakości danych za wystarczającą. Będzie to przedmiotem kolejnych badań.

[illegible]

Tabela 3

Macierz przejścia uzyskana estymatorem Aalena-Nelsona dla okresu (1,5)

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,6434	0,2875	0,0580	0,0043	0,0054	0,0007	0,0000	0,0006
Aa	0,0238	0,6082	0,3000	0,0399	0,0162	0,0035	0,0002	0,0081
A	0,0008	0,0469	0,6916	0,1761	0,0609	0,0175	0,0011	0,0051
Baa	0,0006	0,0181	0,1701	0,5174	0,2099	0,0612	0,0041	0,0186
Ba	0,0025	0,0029	0,0221	0,0939	0,5457	0,2337	0,0198	0,0794
B	0,0004	0,0027	0,0071	0,0186	0,1507	0,4891	0,0482	0,2834
Caa_C	0,0002	0,0009	0,0056	0,0246	0,0660	0,1241	0,1307	0,6479
D	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	1,0000

W celu porównania macierzy należy dokonać ujednolicenia czasu. Oznacza to konieczność wyznaczenia macierzy Aalena dla 1 roku. W tym celu wyznaczony generator macierzy Aalena zlogarytmowano, podzielono przez 4 i eksponowano wykorzystując wzory 5 i 6. Uzyskana macierz przejścia, oznaczona przez AN, została porównana z pozostałymi macierzami. Dodatkowo wyznaczona została średnia arytmetyczna macierzy obliczonych metodą kohort dla okresów jednorocznych, jednakże jej wartości są bardzo bliskie wartościom macierzy PI. Także wartości macierzy wyznaczonej metodą duration okazały się bardzo bliskie wartościom macierzy jednorocznej AN (największa z różnic wyrazów nie przekraczała 0,012). Dlatego macierzy przejścia uzyskanych metodą duration i średnią arytmetyczną nie brano pod uwagę w dalszych rozważaniach.

Tabela 4

Macierz przejścia uzyskana metodą Plug-in (PI)

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,8933	0,1018	0,0036	0	0,0012	0	0	0
Aa	0,0086	0,8781	0,1065	0,0029	0,0019	0	0	0,0019
A	0	0,0153	0,9027	0,0658	0,0126	0,0027	0	0,0009
Baa	0	0,0052	0,0628	0,8356	0,0817	0,0126	0	0,0021
Ba	0,0010	0	0,0029	0,0391	0,8418	0,1035	0,0039	0,0078
B	0	0,0010	0,0019	0,0038	0,0596	0,8221	0,0298	0,0817
Caa_C	0	0	0	0,0138	0,0276	0,0741	0,6034	0,2810
D	0	0	0	0	0	0	0	1,0000

Tabela 5

Macierz przejścia uzyskana metodą GLMM

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,8559	0,0975	0,0276	0	0,0188	0	0	0
Aa	0,038	0,8557	0,072	0	0,0213	0	0	0,0128
A	0	0,0211	0,8943	0,0376	0,0256	0,0147	0	0,0065
Baa	0	0,024	0,0608	0,8019	0,0631	0,0499	0	0
Ba	0,0067	0	0,0121	0,0533	0,7939	0,0627	0,0332	0,0377
B	0	0,0067	0,0167	0,0289	0,0484	0,7731	0,0602	0,0656
Caa_C	0	0	0	0,0413	0,1224	0,0157	0,4235	0,2553
D	0	0	0	0	0	0	0	1

Tabela 6

Macierz przejścia uzyskana za pomocą brzegowego modelu GLMM

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0.8889	0.1015	0.0030	0.0022	0.0011	0.0004	0	0.0025
Aa	0.0085	0.8764	0.1071	0.0035	0.0011	0.0004	0	0.0026
A	0.0004	0.0147	0.9022	0.0740	0.0043	0.0005	0	0.0034
Baa	0.0001	0.0045	0.0530	0.8338	0.0932	0.0046	0.0002	0.0102
Ba	0.0003	0.0003	0.0118	0.0285	0.8362	0.1005	0.0016	0.0204
B	0.0010	0.0009	0.0007	0.0111	0.0624	0.8345	0.0307	0.0584
Caa_C	0.0047	0.0045	0.0035	0.0011	0.0312	0.0723	0.6049	0.2774
D	0	0	0	0	0	0	0	1.0000

Dla obliczonych macierzy zostały wyznaczone generatory według wzoru (6), a następnie po wprowadzeniu poprawki danej wzorem (7) obliczono macierze przejścia wyznaczone za pomocą generatorów. Macierze te pozwalają na dokonanie ratingu w dowolnym momencie.

Tabela 7

Macierz przejścia uzyskana za pomocą generatora dla metody PI (PI G)

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,8907	0,1016	0,0061	0,0002	0,0012	0,0001	0	0,0001
Aa	0,0086	0,8768	0,1065	0,0040	0,0020	0,0002	0	0,0019
A	0,0001	0,0153	0,9026	0,0657	0,0126	0,0027	0,0001	0,0009
Baa	0,0001	0,0052	0,0628	0,8353	0,0815	0,0127	0,0003	0,0021
Ba	0,0010	0,0002	0,0030	0,0390	0,8417	0,1033	0,0040	0,0078
B	0	0,0010	0,0019	0,0039	0,0595	0,8223	0,0295	0,0819
Caa_C	0	0,0001	0,0006	0,0137	0,0275	0,0734	0,6051	0,2795
D	0	0	0	0	0	0	0	1,0000

Tabela 8

Macierz przejścia uzyskana za pomocą generatora dla metody GLMM (GLMM G)

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,8527	0,0971	0,0276	0,0011	0,0188	0,0009	0,0004	0,0011
Aa	0,0378	0,8518	0,0718	0,0022	0,0214	0,0014	0,0005	0,0129
A	0,0006	0,0211	0,8925	0,0374	0,0256	0,0147	0,0011	0,0067
Baa	0,0008	0,0239	0,0604	0,7949	0,0629	0,0497	0,0033	0,0034
Ba	0,0067	0,0015	0,0122	0,0530	0,7930	0,0625	0,0322	0,0382
B	0,0003	0,0068	0,0168	0,0289	0,0489	0,7730	0,0578	0,0666
Caa_C	0,0006	0,0007	0,0025	0,0401	0,1175	0,0159	0,4361	0,2487
D	0	0	0	0	0	0	0	1,0000

Tabela 9

Macierz przejścia uzyskana za pomocą generatora dla jednorocznego estymatora Aalena-Nelsona (AN)

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0,9037	0,0769	0,0174	0,0008	0,0009	0	0	0
Aa	0,0064	0,8949	0,0805	0,0119	0,0041	0,0005	0	0,0013
A	0,0002	0,0125	0,9177	0,0466	0,0171	0,0048	0,0003	0,0005
Baa	0,0001	0,0048	0,0448	0,8724	0,0549	0,0179	0,0012	0,0035
Ba	0,0006	0,0008	0,0061	0,0243	0,8807	0,0606	0,0053	0,0211
B	0,0001	0,0006	0,0018	0,0053	0,0391	0,8657	0,0114	0,0756
Caa_C	0	0,0002	0,0015	0,0061	0,0169	0,0297	0,7801	0,165
D	0	0	0	0	0	0	0	1

7. Porównanie uzyskanych macierzy i wnioski

W pracy [11] przedstawiono przykłady metryk wykorzystywanych do porównywania macierzy przejścia obliczonych różnymi metodami, a także macierzy obliczonych dla różnych okresów obejmujących lata 1981-2002. Spośród wielu zaproponowanych metryk, wyróżniono metrykę opartą na SVD, czyli na średniej z wartości szczególnych macierzy mobilności $\tilde{P} = P - I$.

$$M_{SVD}(P) = \frac{\sum_{i=1}^N \sqrt{\lambda_i(\tilde{P}^T \tilde{P})}}{N} \quad (8)$$

gdzie λ_i są wartościami własnymi macierzy $\tilde{P}^T \tilde{P}$. Metryka oparta na wartościach szczególnych macierzy pozwala mierzyć „siłę migracji” macierzy przejścia. Im wyższa wartość metryki, tym większa mobilność.

Porównania wyznaczonych macierzy, a co za tym idzie porównania skuteczności metody estymacji macierzy, dokonano dwoma sposobami. Wyznaczono wartości własne macierzy przejścia, macierze mobilności oraz ich wartości szczególne i obliczono miary według wzorów przedstawionych w [11].

Tabela 10

Porównanie macierzy przejścia uzyskanych wybranymi metodami

	PI	PIG	GLMM	AN	GLMM G	BGLMM
MP(P) tr	0.1747	0.1751	0.2288	0.1264	0.2294	0.1747
MD(P) det	0.7594	0.7599	0.8682	0.6188	0.8670	0.7588
M2(P)	0.0229	0.0229	0.0279	0.0235	0.0279	0.0229
MSVD(P)	0.2010	0.2013	0.2467	0.3657	0.2467	0.2007

Wyznaczone miary pozwalają tylko częściowo porównać uzyskane macierze. Uzyskane wyniki zawarte w tabeli 10 pozwalają na stwierdzenie, że wyróżniającą się macierzą jest macierz AN. Macierz ta ma najwyższe wartości wyrazów na przekątnej (najniższa miara śladowa MP tr) oraz najwyższą M_{SVD} , czyli ilustruje największą mobilność. Warto zwrócić uwagę także na fakt, że macierz AN oraz macierze uzyskane za pomocą generatorów zawierają mniej elementów zerowych niż pozostałe macierze.

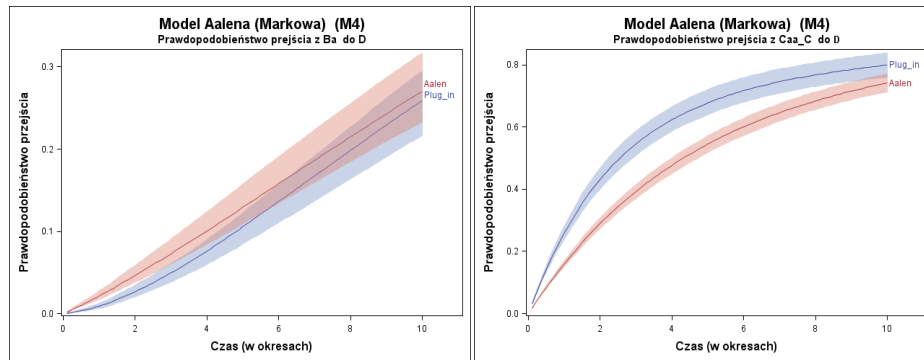
Jednocześnie porównywanie macierzy przejścia przy użyciu „miar odległościowych” nie przekłada się na wynik biznesowy. Z tego powodu proponujemy porównanie oparte na wyznaczeniu dla macierzy przejść generatorów macierzy, a następnie wyznaczenie prawdopodobieństw przejść w funkcji czasu. Dla celów biznesowych krytyczne znaczenia mają prawdopodobieństwa przejść z różnych stanów początkowych do stanu „default”.

W praktyce bankowej wartość instrumentów finansowych wchodzących w skład aktywów banku wyznacza się dyskontując prawdziwe oraz oczekiwane „przepływy pieniężne” przy użyciu rynkowych stóp procentowych. Ryzyko kredytowe wpływa na zmianę czynników dyskontowych poprzez dodanie pewnego przyrostu – spreadu. Wartość spreadu zależy od prawdopodobieństwa defaultu dla danego instrumentu i powinna zmieniać się w zależności od horyzontu czasowego. Stąd naturalną miarą oceny macierzy migracji w praktyce bankowej będzie wartość prawdopodobieństwa przejścia między stanami w funkcji czasu.

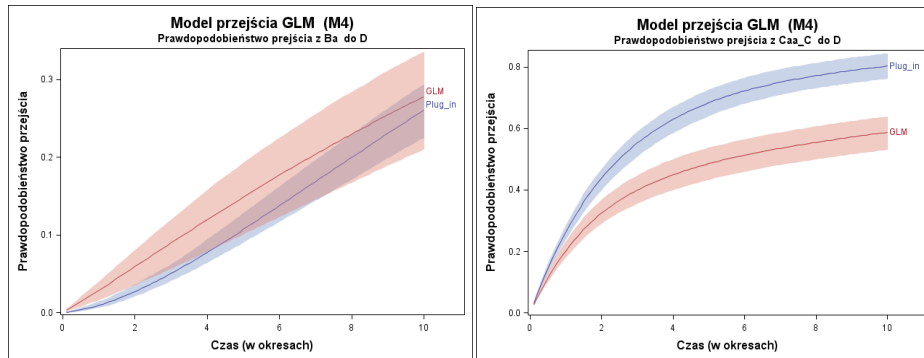
Z uwagi na wykorzystanie danych wygenerowanych sztucznie, prezentujemy porównania krzywych prawdopodobieństw przejść w zależności od horyzontu czasowego oraz testowanie hipotezy równości tych krzywych.

Na rys. 1, 2 przedstawione zostały prawdopodobieństwa przejścia z klas „Ba” i „Caa C” do stanu „default” w funkcji czasu. Oprócz wartości średniej prawdopodobieństw zaznaczone zostały przedziały ufności uzyskane metodą bootstrap.

Poszczególne metody dają inne rezultaty. Na podstawie wykresów można wyciągnąć wniosek, że różne modele mogą być optymalne dla różnych rozkładów. Modele bezpośrednio bazujące na podejściu Markowa nie dopuszczają włączenia zmiennych egzogenicznych i wykorzystują informacje o przejściach na podstawie „bezpośrednich zliczeń”. W praktyce bankowej strategii przyznawania kredytów, a co za tym idzie warunki, do których stosuje się modele, ulegają zmianie, stąd bardzo istotne jest dołączenie dodatkowych czynników do modeli. W ramach modeli statystycznych jest to możliwe poprzez odpowiednie zmienne egzogeniczne.



Rys. 1. Porównanie modelu Aalena i modelu Plug-in dla prawdopodobieństw przejścia z Ba do D oraz z Caa_C do D



Rys. 2. Porównanie modelu GLMM i modelu Plug-in dla prawdopodobieństw przejścia z Ba do D oraz z Caa_c do D

Literatura

1. Agresti A. (2002). *Categorical Data Analysis*. Wiley Series in Probability and Statistics.
2. Basel Committee on Banking Supervision The Internal Ratings-Based Approach. (2001). Consultative Document.
3. Diggle P. (2002). *Analysis of Longitudinal Data*. Oxford University Press, USA.
4. Efron B. and Tibshirani R.J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.
5. Feller W. (1966). *Wstęp do rachunku prawdopodobieństwa*. PWN, Warszawa.
6. Frydman H., Kallberg J.G., Kao L.D. (1985). Testing the Adequacy of Markov Chain and Mover-Stayer Models as Representations of Credit Behavior. *Operations Research*, 33.
7. Frydman H. (1984). Maximum Likelihood Estimation in the Mover-Stayer Model. *Journal of the American Statistical Association*, 79.
8. Höse S. Huschens S. Wania R. (2002). *Rating Migrations. Applied Quantitative Finance: Theory and Computational Tools*, Springer.
9. Iosifescu M. (1987). *Skończone Łańcuchy Markowa*. WNT, Warszawa.
10. Israel R.B., Rosentahl J.S. Wei J.Z. (2001). Finding Generators for Markov Chains via Empirical Transition Matrices, with Applications to Credit Rating. *Mathematical Finance*, 11, 2.
11. Jafry Y., Schuermann T. (2004). Measurement, Estimation and Comparison of Credit Migration Matrices. *Journal of Banking and Finance*, 28, 11.
12. Jarrow R.A., Lando D., Turnbull S.M. (1997). A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies*, 10, 2.
13. Jones M.T. Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk. IMF Working Paper WP/05/219.
14. Lando D., Skodeberg T.M. (2002). Analyzing Rating Transitions and Rating Drift with Continuous Observations. *Journal of Banking and Finance*, 26.
15. Moody's Investors Service. (2000). Historical Default Rates of Corporate Bond Issuers, 1920-1999, 25.
16. Moody's. <http://www.moodys.com/>
17. Rachev S.T. Trueck S. (2009). *Rating Based Modeling of Credit Risk Theory and Application of Migration Matrices*. Academic Press.
18. Saunders A. (2001). *Metody pomiaru ryzyka kredytowego*. Oficyna Ekonomiczna, Kraków.
19. Schuermann T. Credit Migration Matrices w *Encyclopedia Quantitative Risk Analysis & Assessment*. <http://www.wiley.com/legacy/wileychi/risk/>
20. S&P. <http://www.standardandpoors.com/ratings/en/eu>

MODELS OF MIGRATION MATRICES USED IN CREDIT RISK. EXAMPLES AND COMPARISON

Summary

Credit risk models used in banks are based on probability models for occurrence of default. A vast class of the models used in practice (e.g., Credit Metrics) is based on the notion of intensity. In 1997 Jarrow applied Markov chain approach to analyze intensities. The key problem that arises is the selection of appropriate estimators. Within the Markov approach the most frequently used estimators of the migration matrix are cohort and duration estimators. Migration matrices can also be obtained with help of statistical longitudinal models (GLMM) in which states (rating classes) in discrete time points are regarded as matched pairs. In this paper we compare Markov chain models and GLMM models and the influence of their application on bank portfolio evaluation.

Magdalena Rogalska

Politechnika Lubelska

PROGNOZOWANIE CZASU TRWANIA AWARII KOPAREK KOŁOWYCH Z ZASTOSOWANIEM METOD SIECI NEURONOWYCH, REGRESJI WIELORAKIEJ I ARIMA CZĘŚĆ I

Wprowadzenie

Awarye koparek kołowych są zjawiskiem nagminnym. Z analizowanych danych wynika, że w ciągu 10 lat użytkowania jedna koparka wymaga nawet 527 napraw. Powoduje to, że czas wykonania zadania roboczego wydłuża się. Dodatkowy czas potrzebny na naprawy lub możliwość korzystania ze sprzętu zastępczego musi być przewidziany w harmonogramie budowy. Celem pracy jest wyznaczenie dodatkowego czasu potrzebnego na naprawy w określonym okresie użytkowania koparki. Dane do obliczeń uzyskano z bazy sprzętu budowlanego BMTI Sp z.o.o. w Pruszkowie. Wszystkie naprawy z okresu 9 lat są skatalogowane w formie elektronicznej. Znany jest koszt naprawy z podziałem na materiały, sprzęt i robociznę oraz czas naprawy i rodzaj awarii. Do analizy użyto danych dotyczących wszystkich 47 koparek kołowych będących własnością BMTI. Przygotowanie danych do obliczeń wykonano w następujących etapach:

1. Zagregowanie danych w okresy miesięczne – zsumowano dane z każdego miesiąca w odniesieniu do pojedynczych maszyn.

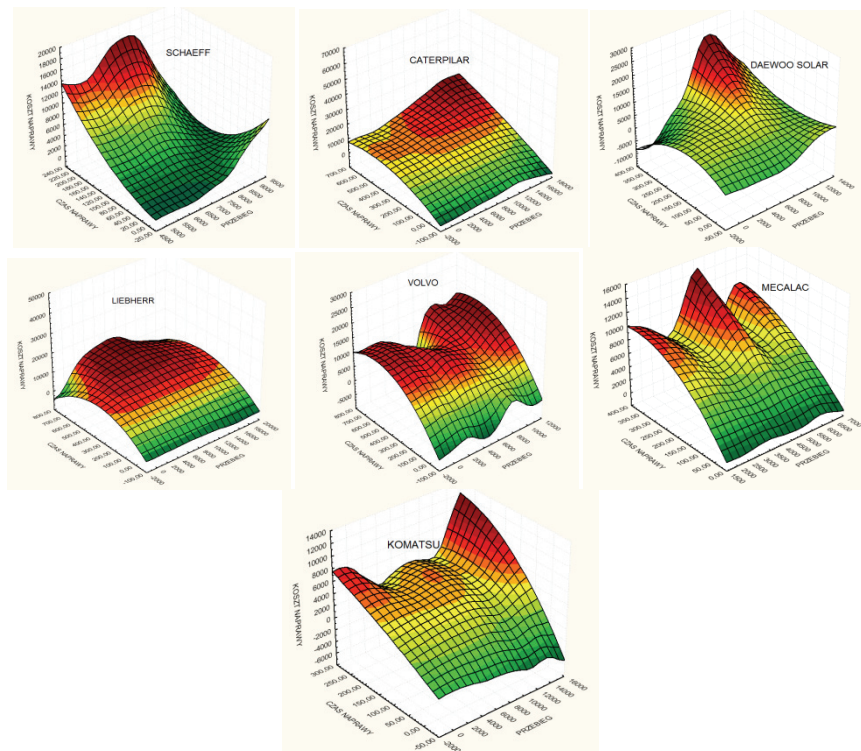
2. Analizowanie danych pod względem możliwości agregacji koparek ze względu na ich markę (analiza sieciami neuronowymi Kohonena, analiza drzew). Wykonano wykresy zależności kosztu i czasu naprawy w zależności od przebiegu maszyn. Wykresy zestawiono na rys. 1. Widać wyraźnie różnice w kształcie wykresów. Zatem istotne jest różnicowanie prognoz pod względem marki koparek.

3. Obliczenia z zastosowaniem zmiennej czas napraw koparki nie dały zadowalających efektów. Występują autokorelacje reszt w metodach regresji wielorakiej i ARIMA. W obliczeniach z użyciem sieci neuronowych jakość uczenia, testowania i walidacji jest w granicach 0,3-0,35.

4. Zmienną czas napraw koparki należy przekształcić w zmienną skumulowany czas napraw koparek. Wówczas zmienna ta posiada trend i zmienność sezonową. Znacznie lepiej nadaje się do prognozowania.

5. Etapem pierwszym obliczeń będzie prawidłowe prognozowanie czasu naprawy jednej koparki. Następnie prognozowanie czasu napraw jednej marki koparek i kolejno poszukiwanie modelu dla wielu marek koparek.

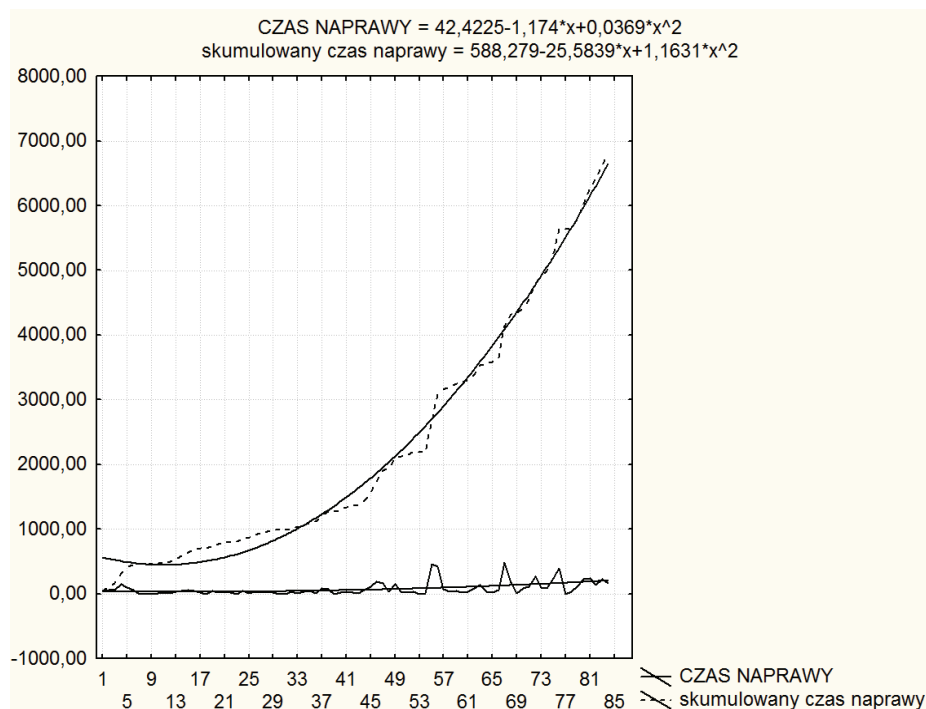
6. W części I pracy należy prawidłowo wyznaczyć modele do prognozowania czasu naprawy jednej koparki. Do analizy wybrano koparkę marki Liebherr L14.



Rys. 1. Wykresy zależności czasu i kosztu napraw koparek kołowych w zależności od ich przebiegu

1. Metoda Regresji Wielorakiej

Do prognozowania czasu napraw koparki kołowej L14 zastosowano metodę regresji wielorakiej. Jako zmienną zależną zastosowano skumulowany czas napraw koparki. Wykres zależności czasu napraw oraz skumulowanego czasu napraw koparki L14 od sierpnia 2001 do lipca 2008 przedstawiono na rys. 2.



Rys. 2. Wykres zależności czasu napraw oraz skumulowanego czasu napraw koparki L14 od sierpnia 2001 do lipca 2008. Na osi x oznaczono okresy odpowiadające kolejnym miesiącom

Wyraźnie widoczny jest trend wielomianowy. Arkusz wyjściowy zawiera zmienne zestawione w tabeli 1. Dekompozycję sezonową zrealizowano metodą $(1,0,-1)$. Metoda polega na tym, że wartość 1 przypisywana jest zmiennej sezonowej wówczas, gdy w zmiennej data jest ten sam miesiąc, np. dla przypadku pierwszego gdy zmienna v_1 to 2001.8 wówczas zmienna v_{43} sierpień przyjmuje wartość 1, a pozostałe zmienne sezonowe (inne miesiące) wartość 0. Przypadki zawierające marzec zawsze dla wszystkich zmiennych v_{36} - v_{47} przyjmują wartość (-1) .

Tabela 1

Zestawienie zmiennych z oznaczeniami do prognozowania skumulowanego czasu napraw koparki kołowej L14

Zmienna	Nazwa zmiennej	Zmienna	Nazwa zmiennej
V1	Data	V9	Czas naprawy
V2	t	V10	Skumulowany czas naprawy
V3	t ²	V11	Opóźniony o 1 okres skumulowany czas naprawy (lag1)
V4	t ³	V12-V35	Opóźniony o odpowiednio od 2 do 25 okresów skumulowany czas naprawy (lag2,...,lag25)
V5	Marka	V36	Styczeń
V6	Numer koparki	V37-V47	Luty,...,grudzień
V7	Wiek	V48	Regresja R1
V8	Przebieg	V49	Reszty RR1=V10-V48

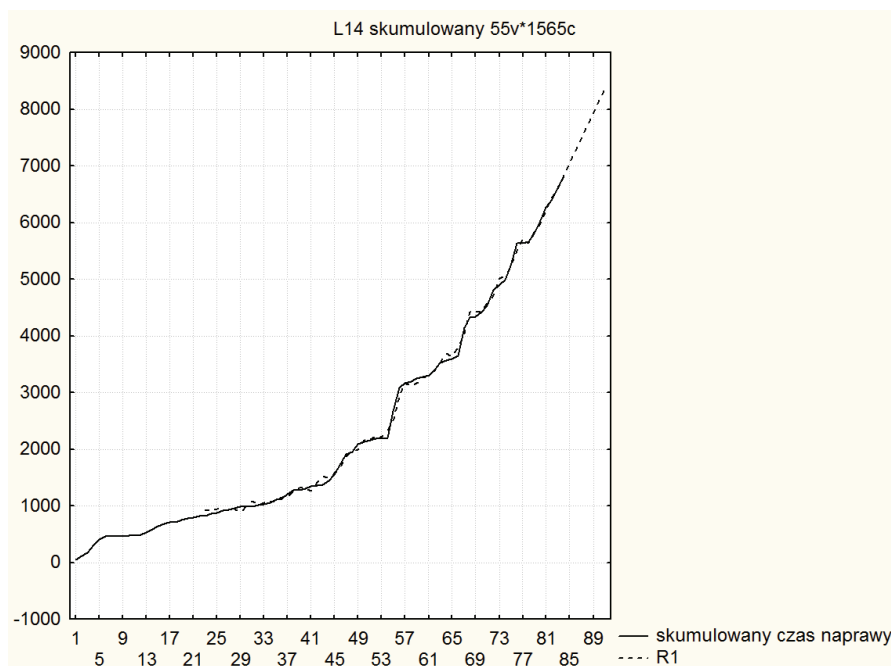
Do obliczeń zastosowano metodę regresji wielorakiej wstecznej. Otrzymano równanie regresyjne w postaci (1), podsumowanie regresji zmiennej zależnej v10 zamieszczono w tabeli 2, wykres zależności zmiennej zależnej v10 i prognozy R1(v10;v2,v3,v11,v12,v32,v37,v47) przedstawiono na rys. 3.

Tabela 2

Podsumowanie regresji model R1(v10;v2,v3,v11,v12,v32,v37,v47)

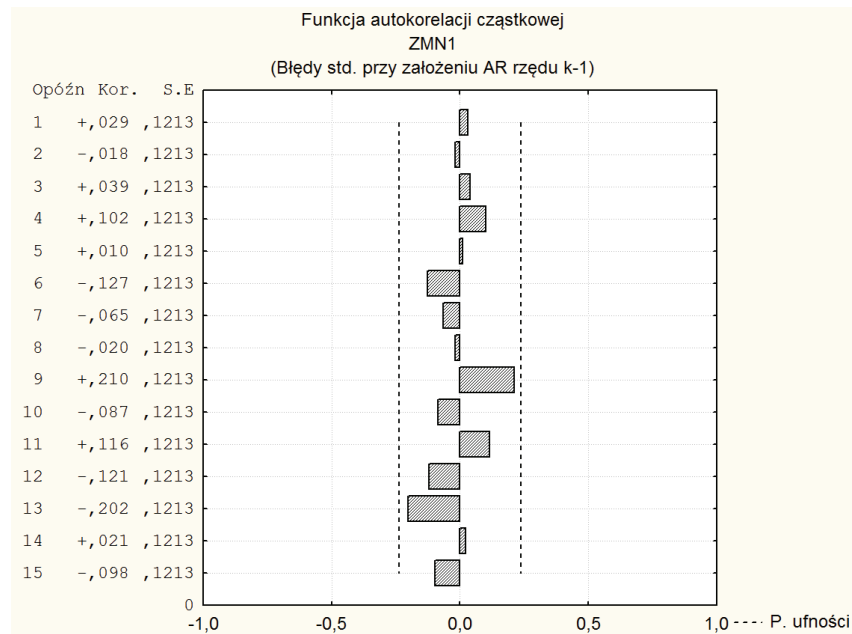
N=62	Podsumowanie regresji zmiennej zależnej: skumulowany czas naprawy (L14 skumulowany) R= ,99917516 R ² = ,99835099 Skoryg. R2= ,99813723 F(7,54)=4670,4 p<0,0000 Błąd standardowy estymacji: 79,086					
	b*	błąd standardowy z b*	b	błąd standardowy z b	t(54)	p
W. wolny			764,8996	185,2680	4,12861	0,000127
t	-0,322996	0,077654	-32,8052	7,8869	-4,15945	0,000115
t ²	0,818822	0,163879	0,7687	0,1538	4,99649	0,000006
lag1	0,978961	0,113035	1,0071	0,1163	8,66072	0,000000
lag2	-0,419328	0,108166	-0,4443	0,1146	-3,87671	0,000289
lag22	-0,064199	0,028930	-0,1300	0,0586	-2,21912	0,030701
m2	0,017942	0,006790	81,2009	30,7280	2,64257	0,010744
m12	-0,020629	0,006525	-93,3594	29,5287	-3,16166	0,002573

$$v_{10} = 906,7561 - 38,3627 \cdot v_2 + 0,8407 \cdot v_3 + 1,0048 \cdot v_{11} - 0,4475 \cdot v_{12} - 0,1718 \cdot v_{32} + 81,7995 \cdot v_{37} - 95,1304 \cdot v_{47} \quad (1)$$



Rys. 3. Wykres zależności skumulowanego czasu naprawy koparki kołowej L14 oraz prognozy uzyskanej metodą regresji $R1(v_{10}; v_2, v_3, v_{11}, v_{12}, v_{32}, v_{37}, v_{47})$

Celem sprawdzenia prawidłowości prognozy metodą regresji wielorakiej wykonano analizę korelacji cząstkowej reszt. Korelogram zobrazowano na rys. 4. Wskazuje on, że nie występuje autokorelacja reszt i model może być przyjęty do dalszych analiz. W tabeli 3 zestawiono wartości prognoz dla okresów od 85 do 90. Poddane będą one analizie błędów w dalszej części pracy.



Rys. 4. Funkcja autokorelacji cząstkowej reszt modelu R1(v10;v2,v3,v11,v12,v32,v37,v47)

Tabela 3

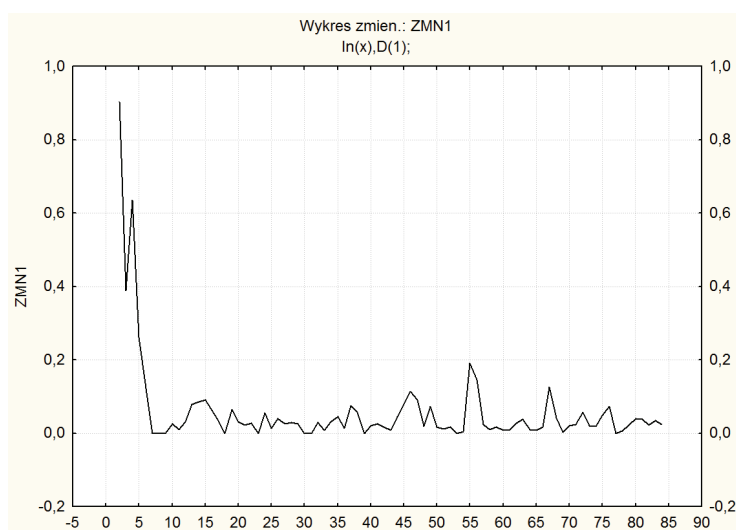
Prognoza modelu R1(v10;v2,v3,v11,v12,v32,v37,v47) skumulowanego czasu
napraw koparki kołowej L14 wartości rzeczywiste

t	Data	Prognoza skumulowanego czasu napraw	Wartości rzeczywiste Skumulowanego czasu napraw
85	2008.08	6978,07	7054,30
86	2008.09	7257,22	7317,63
87	2008.10	7511,64	7589,30
88	2008.11	7764,70	7700,96
89	2008.12	7686,67	7850,96
90	2009.01	7963,75	8119,30

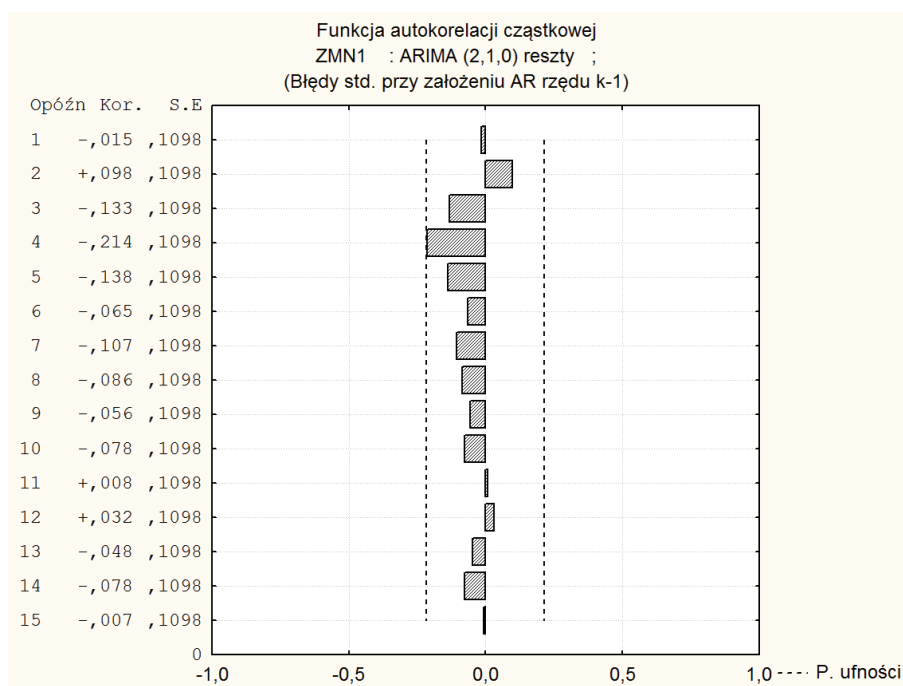
W równaniu regresji wielorakiej występuje zmienna v11, opóźniony o 1 okres skumulowany czas naprawy (lag1). Oznacza to, że przy wykorzystaniu tego modelu można prognozować czas napraw z wyprzedzeniem o 1 okres – tylko na 1 miesiąc do przodu. Model wymaga przygotowania arkusza kalkulacyjnego zawierającego 47 przekształconych danych, jest bardzo pracochłonny.

2. Metoda ARIMA

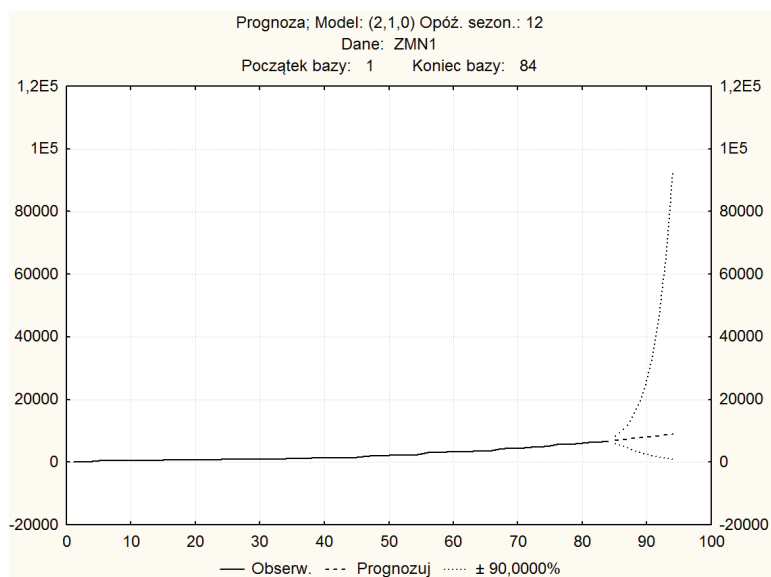
Szereg czasowy skumulowany czas napraw koparek kołowych poddano analizie ARIMA. Celem jest wyznaczenie prognozy na okresy od sierpnia 2008 do stycznia 2009 roku. Prawidłowość wyznaczenia prognozy sprawdzana jest poprzez funkcję autokorelacji cząstkowych reszt, jeśli nie ma autokorelacji reszt to model nadaje się do przyjęcia. W wyniku analizy wielu modeli przyjęto model ARIMA (2,1,0) jednokrotnie różnicowany i zlogarytmowany. Szereg czasowy po różnicowaniu i zlogarytmowaniu przedstawiono na rys. 5. Funkcję autokorelacji cząstkowej reszt modelu ARIMA (2,1,0) zobrazowano na rys. 6. Wartości prognozowane skumulowanego czasu napraw koparki L14 na okresy 85-90 zestawiono w tabeli 4. Wykres danych i prognozy ARIMA (2,1,0) dla szeregu przekształconego przedstawiono na rys. 7.



Rys. 5. Szereg czasowy skumulowany czas napraw koparki L14 po jednokrotnym różnicowaniu i zlogarytmowaniu



Rys. 6. Funkcja autokorelacji cząstkowej reszt modelu ARIMA (2,1,0)



Rys. 7. Wykres danych i prognozy ARIMA (2,1,0) szereg przekształcony

Tabela 4

Prognoza model ARIMA (2,1,0)

Nr obserwacji	Prognoza; Model: (2,1,0) Opóźnienie sezon.: 12 Dane: ZMN1 Początek bazy: 1 Koniec bazy: 84		
	prognoza	dolne 90,0000%	górne 90,0000%
85	7002,219	5945,762	8246,39
86	7193,568	5309,573	9746,06
87	7396,446	4550,231	12023,00
88	7602,225	3815,607	15146,69
89	7814,992	3123,707	19551,80
90	8033,149	2506,672	25743,89

Istnieje możliwość prognozowania skumulowanego czasu napraw koparek z wykorzystaniem metody ARIMA. W proponowanym modelu nie występuje autokorelacja reszt, zatem model może być przyjęty do dalszych analiz. Metoda ARIMA nie wymaga zastosowania predyktorów innych niż skumulowany czas napraw. Jest najłatwiejszą metodą z wszystkich możliwych, o ile zostaną wcześniej opanowane metody wyznaczania parametrów. Wyniki obliczeń uzyskanych metodą ARIMA zostaną poddane analizie błędów w 4 rozdziale pracy.

3. Sieci neuronowe

Analizowano wiele możliwości przyjmowania zmiennych wejściowych z arkusza kalkulacyjnego regresji wielorakiej. Zasadą podstawową jest przyjęcie modelu jak najbardziej prostego, spełniającego postulat maksimum jakości uczenia, testowania i walidacji. Modelami takimi okazały się modele, w których danymi wejściowymi były t , t^2 , t^3 oraz planowany przebieg maszyn. Zastosowano sieci MLP (nie znaleziono podstawy do stosowania sieci RBF). Podsumowanie aktywnych sieci MLP zestawiono w tabeli 5, a przewidywania dla danych od sierpnia 2008 do stycznia 2009 roku dla sieci MLP w tabeli 6.

Tabela 5

Podsumowanie aktywnych sieci MLP

Id sieci	Nazwa sieci	Jakość (uczenie)	Jakość (testowanie)	Jakość (walidacja)
1	MLP 4-3-1	0,996851	0,998577	0,999205
2	MLP 4-3-1	0,996859	0,999091	0,998913
3	MLP 4-3-1	0,998019	0,999466	0,998989
4	MLP 4-3-1	0,996901	0,998452	0,998939
5	MLP 4-4-1	0,996619	0,998498	0,999080
	Błąd (uczenie)	Błąd (testowanie)	Błąd (walidacja)	
1	10505,30	5067,755	5351,535	
2	10536,76	3063,102	6362,606	
3	6473,42	2278,636	5321,129	
4	10119,00	5195,530	5876,203	
5	11255,67	4661,459	4577,456	
	Algorytm uczenia	Funkcja błędu		Aktywacja (ukryte)
1	BFGS 14	SOS		Liniowa
2	BFGS 18	SOS		Tanh
3	BFGS 71	SOS		Tanh
4	BFGS 22	SOS		Liniowa
5	BFGS 12	SOS		Liniowa

Tabela 6

Przewidywania dla danych od sierpnia 2008 do stycznia 2009 roku
dla sieci MLP

Nr obserwacji	Przewidywania dla nowych danych (L14 skumulowany)				
	1.MLP 4-3-1	2.MLP 4-3-1	3.MLP 4-3-1	4.MLP 4-3-1	5.MLP 4-4-1
85	7017,745	6904,613	6914,752	7106,912	7035,355
86	7186,234	7069,923	7072,172	7297,861	7234,405
87	7356,023	7239,354	7235,043	7497,974	7439,126
88	7539,247	7422,733	7426,669	7733,309	7657,967
89	7712,691	7603,418	7598,918	7954,689	7875,057
90	7897,011	7797,953	7790,147	8205,764	8104,530

cd. tabeli 6

Nr obserwacji	Przewidywania dla nowych danych (L14 skumulowany)			
	t	t ²	t ³	PRZEBIEG
85	85,00000	7225,000	614125,0	18149,00
86	86,00000	7396,000	636056,0	18387,00
87	87,00000	7569,000	658503,0	18598,00
88	88,00000	7744,000	681472,0	18643,00
89	89,00000	7921,000	704969,0	18785,00
90	90,00000	8100,000	729000,0	18791,00

Istnieje możliwość prognozowania skumulowanego czasu napraw koparek z wykorzystaniem metody automatycznych sieci neuronowych. Jakość uczenia, testowania i walidacji osiąga wartości około 0,99. Daje to wysokie prawdopodobieństwo uzyskania wyniku MAPE mniejszego niż 5%, czyli bardzo dobrego. Danymi wejściowymi w proponowanych modelach są kolejne okresy i przewidywany przebieg maszyn (wartość znana z kosztorysu). Zatem istnieje możliwość prognozowania na wiele okresów do przodu. Wyniki uzyskanych obliczeń zostaną poddane analizie błędów w 4 rozdziale pracy.

4. Analiza wyników

W tabeli 7 zebrano wyniki obliczeń metodami regresji, ARIMA i sieci neuronowych. Na rys. 8 przedstawiono interpretację graficzną wyników. Będą one porównywane z wartościami liczbowymi rzeczywistymi, skumulowanego czasu napraw koparki L14, uzyskanymi w okresach od 85 do 90. Obliczane będą kolejno błędy dla różnych okresów (od jednego do sześciu). W ten sposób można ocenić na ile okresów w przód można postawić właściwą prognozę w analizowanym przypadku. Obliczone błędy zestawiono w tabeli 8, błędy MAPE zilustrowano dla poszczególnych prognozowanych okresów na rys. 9.

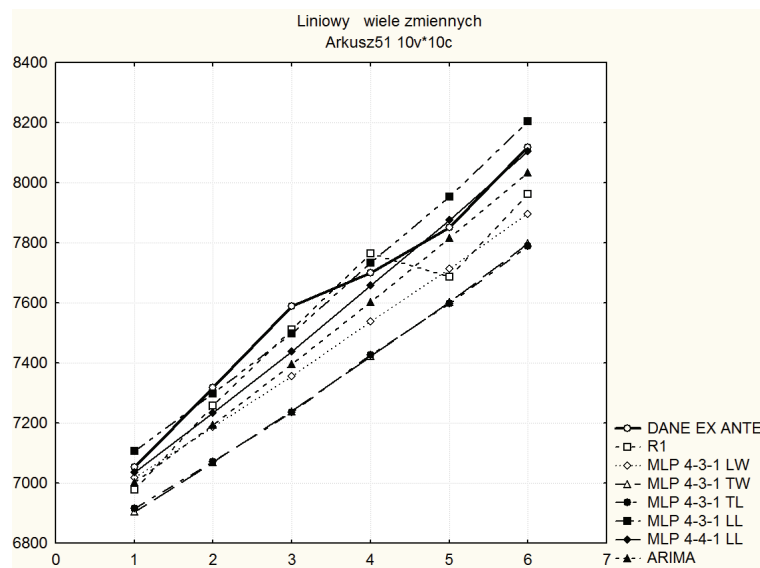
Tabela 7

Zestawienie wyników wartości prognozowanych i uzyskanych danych dla kolejnych okresów

Nr obserwacji	DANE	REGRESJA R1	1.MLP 4-3-1	2.MLP 4-3-1
85	7054,303	6978,0714	7017,745	6904,613
86	7317,636	7257,224	7186,234	7069,923
87	7589,303	7511,6403	7356,023	7239,354
88	7700,97	7764,7057	7539,247	7422,733
89	7850,97	7686,6724	7712,691	7603,418
90	8119,303	7963,7506	7897,011	7797,953

cd. tabeli 7

Nr obserwacji	3.MLP 4-3-1	4.MLP 4-3-1	5.MLP 4-4-1	ARIMA(2,1,0)
85	6914,752	7106,912	7035,355	7002,219
86	7072,172	7297,861	7234,405	7193,568
87	7235,043	7497,974	7439,126	7396,446
88	7426,669	7733,309	7657,967	7602,225
89	7598,918	7954,689	7875,057	7814,992
90	7790,147	8205,764	8104,530	8033,149



Rys. 8. Prognozowane wartości skumulowanego czasu awarii koparki L14 od sierpnia 2008 do stycznia 2009 w zestawieniu z wartościami rzeczywistymi (DANE)

Średni absolutny błąd procentowy obliczany jest zgodnie ze wzorem 2, wyniki obliczeń zestawiono w tabeli 8 i przedstawiono na rys. 9.

$$MAPE = \frac{1}{T-n} \sum_{i=T-n}^T \frac{|Y_i - Y_p|}{Y} \quad (2)$$

gdzie:

MAPE – średni absolutny procentowy błąd (mean absolute percentage error),

T – suma ilości okresów obliczeniowych i prognozowanych,

n – ilość okresów prognozowanych,

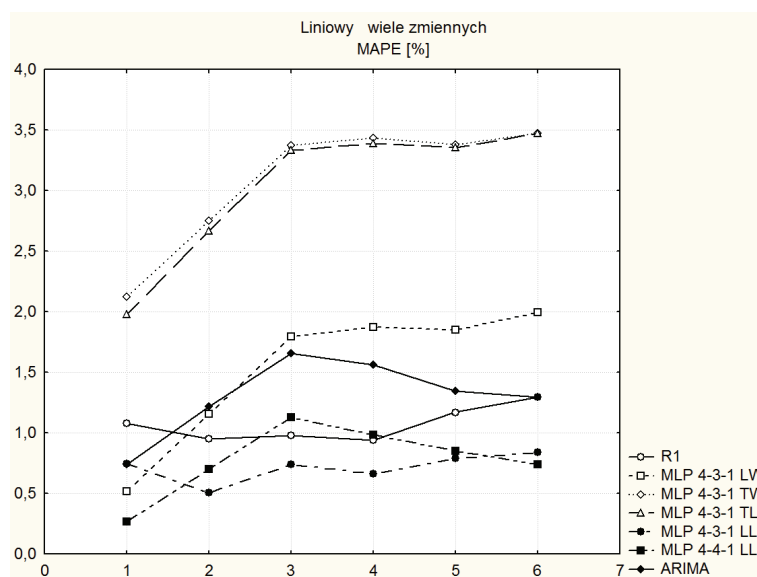
Y_i – wartość rzeczywista zmiennej w okresie i,

Y_p – wartość prognozowana zmiennej w okresie i.

Tabela 8

Zestawienie wyników błędu MAPE wartości prognozowanych i uzyskanych danych dla kolejnych okresów

Nr obserwacji	REGRESJA R1	1.MLP 4-3-1	2.MLP 4-3-1	
85	1,080642	0,518235	2,121966	
85-86	0,953108	1,156967	2,753563	
85-87	0,976512	1,795910	3,372735	
85-88	0,939293	1,871939	3,432803	
85-89	1,169974	1,849809	3,376870	
85-90	1,294285	1,997811	3,473701	
Nr obserwacji	3.MLP 4-3-1	4.MLP 4-3-1	5.MLP 4-4-1	ARIMA(2,1,0)
85	1,978239	0,745773	0,268603	0,738337
85-86	2,666331	0,508010	0,703009	1,216907
85-87	3,333518	0,739803	1,128273	1,658329
85-88	3,390614	0,659836	0,985806	1,564308
85-89	3,354581	0,792090	0,850005	1,343099
85-90	3,471150	0,837555	0,738664	1,296100



Rys. 9. Błędy MAPE dla prognozowanych wartości skumulowanego czasu awarii koparki L14 od sierpnia 2008 do stycznia 2009

Wszystkie zaproponowane modele spełniają wymagania bardzo dobrego prognozowania. Dla wszystkich modeli średni absolutny błąd procentowy $MAPE < 5\%$. Najlepszy wynik uzyskano dla modelu sieci neuronowych MLP 4-3-1 LL, $MAPE < 1\%$.

5. Obliczenie opóźnień

Obliczenia prowadzone były dla zmiennej skumulowany czas napraw koparki. Należy powrócić do zmiennej wyjściowej czas napraw. Obliczone wartości zestawiono w tabeli 9.

Tabela 9

Zestawienie wyników wartości prognozowanych i uzyskanych danych dla kolejnych okresów [r-g]

DATA	DANE	REGRESJA R1	1.MLP 4-3-1	2.MLP 4-3-1
09.2008	238	279	168	165
10.2008	211	254	169	169
11.2008	45	253	183	183
12.2008	142	-78	173	180
01.2009	6	277	184	194
DATA	3.MLP 4-3-1	4.MLP 4-3-1	5.MLP 4-4-1	ARIMA(2,1,0)
09.2008	157	190	199	191
10.2008	162	200	204	202
11.2008	191	235	218	205
12.2008	172	221	217	212
01.2009	191	251	229	218

W metodzie regresji wielorakiej 12.2008 otrzymano wynik -78 . Jest to wartość nieprawidłowa. Pomimo poprawności statystycznej model nie nadaje się do przyjęcia, nie spełnia wymagań racjonalności. Model zostaje odrzucony na ostatnim etapie obliczeń.

Otrzymane wartości czasu napraw koparki są bardzo wysokie. Jest to 10 rok jej użytkowania. Należało się spodziewać tak długich przestojów potrzebnych na naprawy.

Podsumowanie

W harmonogramowaniu procesów budowlanych bardzo istotne jest wyznaczanie zapasów czasu. Tworzone są bufony czasu wyznaczane standardowo, np. Goldratt wyznacza bufor projektowy, który zawsze ma wartość 25% pierwotnie ustalonego czasu trwania procesu. Znajomość rzeczywistych wartości buforów czasu jest bardzo istotna. Mając rzeczywiste dane proces decyzyjny może osiągnąć wyższy poziom prawidłowości planowania. Dotychczas w budownictwie nie używano dużych baz danych. Jest to jednak możliwe, co wykazano w niniejszej pracy. Dane coraz częściej zbierane są przez firmy budowlane oraz przez firmy związane z budownictwem.

Kontynuacją niniejszej pracy będzie prognozowanie czasu awarii koparek kołowych z podziałem na marki.

Literatura

1. Cieślak M. (2001). Prognozowanie gospodarcze: metody i zastosowanie. Wydawnictwo Naukowe PWN, Warszawa.
2. Kot S., Jakubowski J., Sokołowski A. (2007). Statystyka. Difin, Warszawa.
3. Martyniak Z. (1997). Wstęp do inwentyki. Wydawnictwo Uczelniane, Kraków, s. 32.
4. Podręcznik internetowy STATISTICA. <http://www.statsoft.pl/textbook/sttimser.htm>
5. Skorupka D. (2004). Neural Networks in Risk Management of a Project. 2004 AACE International Transaction (CSC.1.51-CSC.1.57). The Association for the Advancement of Cost Engineering, USA, Washington.
6. Skorupka D. (2003). Risk Management in Building Projects. AACE International Transaction (CSC.1.91-CSC.1.96). The Association for the Advancement of Cost Engineering, USA, Orlando.
7. Stanisław A. (2006). Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. T. 1. StatSoft Polska Sp. z o.o., Kraków.

**PREDICTING THE DURATION OF FAILURE WHEEL EXCAVATORS USING
NEURAL NETWORKS, MULTIPLE REGRESSION AND ARIMA METHODS.**

PART 1

Summary

The study analyzed the possibility of forecasting the time of failure wheeled excavators using the methods of multiple regression, ARIMA (Autoregressive Integrated Moving Average) and automated neural network. There were presented : the methodology for the preparation of data sheets, transformations of variables and selection of models in terms of statistical regularity of their use. The calculation results of forecasts were analyzed for periods from 85 to 90 by calculating the mean absolute percentage error (MAPE). Work presented the advantages and disadvantages of the proposed methods. The presented methodology is a new solution, not published, yet. There are plans to develop a method of generalization of computations.

Stanisław Walukiewicz

Instytut Badań Systemowych PAN w Warszawie

TWIERDZENIE PITAGORASA W NAUKACH SPOŁECZNYCH

Wprowadzenie

Powszechnie wiadomo, że w gospodarce rynkowej w stanie równowagi ekonomicznej prawdziwa jest równość

$$\text{popyt} = \text{podaż} \quad (1)$$

Jest to przykład równości tu i teraz, gdyż obecnie nie jest ona prawdziwa np. na Kubie czy w Korei Północnej, a w czasach tzw. komuny nie obowiązywała w Polsce, gdy mieliśmy kartki na mięso i cukier. Zatem znak = ma znacznie węższe znaczenie w (1) niż np. w równości $2 + 3 = 5$, która jest prawdziwa zawsze i wszędzie w naszym realnie istniejącym świecie.

Zasadniczym celem opracowania jest pokazanie, że chociaż (1) jest powszechnie znane i opisane w tysiącach, a być może w milionach książek, artykułów i publikacji, to potraktowanie go jako równości (typu) tu i teraz pozwala sformułować kilka oryginalnych wniosków i uwag. Ściśle rzecz biorąc, (1) mówi, że zagregowana wartość popytu równa się tu teraz (zagregowanej) wartości podaży. Termin „wartość”, podobnie jak „rynek” odgrywa w naszych badaniach bardzo istotną rolę, dlatego w referacie podamy nowe i bardzo ogólne definicje obu tych pojęć. Następnie sformułujemy zasadę ortogonalności, która prezentowaliśmy na poprzedniej konferencji [9]. W konkluzji sformułujemy odpowiednik twierdzenia Pitagorasa w naukach społecznych, takich jak ekonomia, zarządzanie, socjologia, psychologia czy nauki polityczne.

1. Wartość i rynek

Wszyscy codziennie używamy takich pojęć jak „wartość” oraz „rynek” – i co więcej – pojęcia te mają coraz szersze znaczenie. Oprócz wartości materialnych i niematerialnych mówimy często o wartościach moralnych, artys-

tycznych, wartości życia ludzkiego itd. Podobnie, prócz powszechnie znanego rynku towarów i usług, mamy rynek emocji (np. sportowych), rynek bankowy, rynek firm sprzedawanych za symboliczną złotówkę, rynek wewnętrzny (wewnątrz danej firmy) itd. Zatem potrzebne nam są bardzo ogólne, szerokie definicje tych pojęć.

Termin „wartość” jest powszechnie używany i ma bardzo głęboki filozoficzny sens, podobnie jak pojęcie czasu czy jakości. Prawie każdy z nas miałby trudności w udzieleniu zwięzłej odpowiedzi na pytanie, co to jest „czas” lub „jakość”, natomiast każdy doskonale wie, co to jest brak czasu czy brak jakości. O ile w przypadku czasu i jakości taka wiedza zwykle wystarcza w codziennej praktyce, o tyle określenie wartości jako tego, co nie jest bezwartościowe, to tautologia, klasyczne „masło maślane”, które nie posuwa ani o krok naprzód naszych rozważań.

Definicja 1

Wartość to ekonomiczny ekwiwalent (równoważnik) wszystkiego, zwykle wyrażany w jednostkach monetarnych.

Jest to definicja typu tu i teraz. Mówi ona, że wszystko na tym świecie ma swoją wartość, swój ekonomiczny ekwiwalent, który jest wyznaczany tu i teraz i zwykle mierzony w pieniądzu. Słowo „zwykle” w poprzednim zdaniu należy rozumieć tak, że np. wartość absolwenta uczelni mierzy się ocenami w jego indeksie i na dyplomie, które są/powinny być uwzględniane w procedurze przyjmowania tego absolwenta do pracy. Zwykle w takiej procedurze wycenia się wartość kandydata w punktach, które później przelicza się na pieniądze (jego wynagrodzenie), co, jak pokażemy w następnym punkcie, jest wartością kapitału ludzkiego tej osoby na rynku wewnętrznym rozważanej firmy.

Definicję 1 można interpretować w następujący sposób. Każdy z nas doskonale wie, jak działa (klasyczna) waga szalkowa. Na przykład na jednej szalce kładziemy jabłko a na drugiej odpowiednie odważniki tak, aby waga była w równowadze. Wyobraźmy sobie teraz szalkową wagę wartości, która działa bardzo podobnie. Jeśli dane jabłko kosztuje tu i teraz złotówkę, to szalkowa waga wartości będzie w równowadze, gdy na jednej szalce położymy to jabłko a na drugiej jednozłotową monetę lub dwie pięćdziesięciogroszówki lub pięć monet 20-groszowych itd. Zatem jeden złoty jest tu i teraz (ekonomicznym) równoważnikiem, tj. wartością danego jabłka. Zauważmy, że zwykle złotówka jest lżejsza niż typowe jabłko, a więc klasyczna waga szalkowa w tym przypadku nie będzie w fizycznej równowadze.

Z definicji 1 wynika, że dowolny materialny lub niematerialny element naszego świata, np. rzecz, idea (patent), talent piłkarza zawodowego itp. ma swoją wartość, swój ekwiwalent ekonomiczny mierzony tu i teraz w pieniądzu, który może się różnić od jego wartości określonej tam i wtedy. By uniknąć ewentualnych nieporozumień: autor nigdy nie twierdził i nie twierdzi, że wszystkie ekwiwalenty ekonomiczne istniejące tu i teraz są doskonałe i nie powinny być nigdy zmieniane. Podręcznik [10] daje teoretyczną podbudowę i praktyczne rady dla takich zmian, o ile są one potrzebne.

Jak wiemy, w gospodarce rynkowej z prawa popytu i podaży (1) wynika wartość życia ludzkiego (np. pasażera samolotu), wartość talentu piłkarza, wartość kompetencji, doświadczenia itp., np. prezesa wielkiej korporacji czy wartość atrakcyjnego wyglądu modelki. Zauważmy jeszcze, że wartości artystyczne (obrazu, rzeźby, utworu muzycznego itp.) są wyceniane na odpowiednio ukształtowanych (zdefiniowanych) rynkach. Podobnie wartości moralne danego autora czy filozofa są lub powinny być wyceniane tu i teraz, zgodnie z (1), poprzez jego wynagrodzenia, nakłady/sprzedaż książek czy honoraria za wykłady, odczyty itp. Możemy zatem podsumować nasze rozważania w formie bardzo ważnego wniosku.

Wniosek 1

Każda wartość jest wyznaczana zawsze, tylko i wyłącznie na odpowiednio zdefiniowanym realnym lub umownym rynku.

Inaczej rzecz ujmując, wartość to liczba jednostek monetarnych, które ktoś, jako wolny człowiek, gotów jest zapłacić w rzeczywistej lub umownej transakcji na odpowiednio zdefiniowanym rynku. Z wniosku 1 wynika, że w gospodarce rynkowej wartości czegokolwiek nie można określać poza realnym lub umownym rynkiem, który tu i teraz jest powszechnie uważany za odpowiednio zdefiniowany. Wniosek ten mówi, że wartość jest zawsze, a nie tu i teraz, definiowana na rzeczywistym lub umownym rynku. Zatem potrzebna jest nam bardzo ogólna definicja rynku, obejmująca rzeczywiście istniejące rynki, jak też rynki umowne, będące np. owocem teoretycznych rozważań.

Definicja 2

Rynek to reguły gry, które tu i teraz są powszechnie uważane za uczciwe, sprawiedliwe i adekwatne. Te reguły mówią, jak należy rozumieć prawo (1) o równowadze między popytem i podażą.

Jest to bardzo ważna definicja, która, podobnie jak definicja 1, po raz pierwszy została podana w podręczniku [10] i dlatego w następnym punkcie na wielu przykładach wyjaśnimy, jak należy ją rozumieć. Przykłady te dowodzą

tezy, że rynek jest wszędzie, że reguły gry rynkowej zmieniają się w czasie, że to co dokładnie opisuje rzeczywistość tu i teraz, nie musi być takim tam i wtedy. Między innymi, na przykładzie rynku emocji związanych z piłką nożną wyjaśnimy, dlaczego, naszym zdaniem, reguł gry rynkowej, uważanych za uczciwe i sprawiedliwe w Wielkiej Brytanii, nie udało się wprowadzić w Polsce.

2. Przykłady rynków

W tym punkcie korzystając z definicji 2 opiszemy rynki zarówno istniejące od stuleci, jak też te wynalezione (zdefiniowane) w ostatnich 30-40 latach. Świadomie używamy zwrotu „wynaleźć rynek”, gdyż jak się Czytelnik przekona właśnie wynalezienie reguł gry rynkowej, które tu teraz są powszechnie uważane za uczciwe, sprawiedliwe i odpowiednie, nie jest sprawą łatwą.

2.1. Powszechny rynek towarów i usług

Powszechny rynek towarów i usług ma reguły ukształtowane przez tysiącletnią praktykę, niemal od zarania ludzkości. Zatem organizatorem powszechnego rynku towarów i usług jest tradycja i zdrowy rozsądek. Te reguły w powszechnej opinii tu i teraz są uważane za uczciwe, sprawiedliwe i odpowiednie, a wszelkie odstępstwa od nich, np. oszustwa, korupcja itp., są wyjątkami, które je tylko dodatkowy raz potwierdzają. Rynek ten w początkowym stadium rozwoju był pewnym zgrupowaniem jarmarcznych (bazarowych) stoisk i właśnie jarmark, targ, bazar itp. weźmiemy jako punkt startu naszej analizy.

Tak rozumiany rynek tworzą aktorzy rynku, którzy dzielą się na sprzedawców lub aktorów (rynku) po stronie podaży i kupujących – aktorów (rynku) po stronie popytu. Ci ostatni mogą być aktywnymi uczestnikami rynku, tj. kupować lub też być pasywnymi jego obserwatorami czyli robić, mówiąc z angielska, window shopping. Natomiast aktorzy rynku po stronie podaży pełnią na nim aktywną rolę (sprzedają, reklamują, atrakcyjnie prezentują swoje towary/usługi itp.). Rynek jest rzeczywistym realnym rynkiem wtedy i tylko wtedy, kiedy istnieją na nim te dwie grupy aktorów. Zatem rynek bez potencjalnych/rzeczywistych kupujących nie jest rynkiem, a wystawą, która poniosła klęskę.

Te dwie grupy aktorów rozmawiają, a więc nawiązują w naszej terminologii relacje, zwykle relacje nieformalne [10]. Zatem niezwykle trudno wyobrazić sytuację, aby aktorem po stronie podaży była osoba głuchoniema. Przykład dawnego Stadionu Dziesięciolecia, nazwanego słusznie największym jarmarkiem Europy, gdzie do 2008 roku handlowali ludzie dosłownie z całego

świata, pokazuje, że ci aktorzy nie muszą wcale znać polskiego ani żadnych języków obcych. Jest to jeszcze jeden dowód na siłę przyciągania idei powszechnego rynku, która bez gromkich deklaracji, w codziennym trudzie dowodzi, że wszyscy ludzie są równi (na rynku).

Różny jest próg wejścia na rynek dla tych dwóch grup aktorów rynkowych. Zwykle wstęp na jarmark, targ itp. jest bezpłatny dla kupujących, czyli jest zerowy próg wejścia na rynek dla aktorów po stronie popytu. Natomiast sprzedawcy muszą wnieść pewne opłaty (placowe, opłaty za wynajęcie stoiska itp.). Zatem po stronie podaży mamy pewien, niezerowy próg wejścia na rynek. Zauważmy, że zerowy próg wejścia na rynek po stronie popytu wynika wprost z jego powszechności, z faktu, że tylko powszechny czyli popularny rynek gwarantuje rzeczywistych nabywców, bo aktywni aktorzy rynku po stronie popytu nie spadają z nieba, a biorą się tylko i wyłącznie z potencjalnych kupujących. Co więcej, powszechność rynku wymusza uczciwość jego reguł, gdyż sprzedawca wykonujący dziennie, powiedzmy, kilkadziesiąt transakcji, ryzykuje zbyt wiele oszukując jednego lub kilku klientów.

Reguły powszechnego rynku towarów i usług nie tylko mówią, jak należy rozumieć tu i teraz równowagę między popytem i podażą, czyli równanie (1), ale również zmuszają sprzedawcę do podawania uczciwej (prawdziwej) ceny potencjalnemu nabywcy. Z powszechności rynku wynika, że nabywca ma informację o aktualnych (tu i teraz) cenach i jest wielu sprzedawców na rynku. Jeśli dany sprzedawca poda zbyt wysoką cenę, to nabywca pójdzie do innego sprzedawcy. I odwrotnie, jeśli sprzedawca poda zbyt niską cenę, to nabywca z potencjalnego stanie się rzeczywistym i sprzedawca straci. Zatem w każdej sytuacji nieuczciwy sprzedawca jest karany i to reguły powszechnego rynku, a nie żadna religia czy zasady dobrego wychowania, wymuszają uczciwe zachowanie sprzedawcy w stosunku do potencjalnego nabywcy.

Warto zauważyć, że te reguły nie są zwykle nigdzie spisane, ani nie są zatwierdzane przez tzw. wyżej stojące organy. Co więcej, historia rozwoju gospodarczego dobitnie pokazuje, że administracyjne próby „poprawiania” reguł powszechnego rynku przynosiły zwykle więcej szkody niż pożytku. W istocie reguły te bazują na dwóch bardzo ogólnych założeniach.

Założenie 1. Podstawy powszechnego rynku towarów i usług

- Z1. Wszyscy aktorzy zarówno po stronie popytu, jak i podaży mają równy dostęp do informacji o rynku (ceny, warunki sprzedaży itp.).
- Z2. Próg wejścia na rynek dla wszystkich aktorów po stronie podaży jest taki sam, a dla aktorów po stronie popytu jest równy zero (nie istnieje).

Założenia te mówią, że na powszechnym rynku towarów i usług wszyscy są równi, jeżeli chodzi o dostęp do informacji i wstęp na ten rynek, czyli o rzeczywisty lub potencjalny udział w grze rynkowej.

W miarę rozwoju społeczno-gospodarczego powstają coraz to nowe rynki z coraz wyższymi progami wejścia. Na przykład, w naszej kulturze nie kupuje się nowych samochodów w jarmarcznych budach, a tylko u autoryzowanych sprzedawców, którzy mogli przekroczyć próg wejścia na rynek nowych samochodów (odpowiednie inwestycje, przeszkolenie, wiarygodność finansowa itp.). Zatem rynki wyrobów/usług wysokiej technologii (rynek samochodowy, rynek komputerów i oprogramowania, rynek usług bankowych itp.) charakteryzują się wysokimi progami wejścia dla aktorów po stronie podaży. Jest też rzeczą oczywistą, że w miarę powstawania rynków, na których handluje się coraz bardziej skomplikowanymi towarami/usługami, pojawiają się odstępstwa od założenia Z1 oraz Z2. Na przykład USA wprowadziły embargo na eksport wyrobów wysokiej technologii do niektórych krajów (Iran, Kuba, Korea Północna itd.). Wtedy dla tych krajów próg wejścia po stronie popytu na amerykański rynek wyrobów wysokich technologii jest nieskończenie wysoki. Podobnie jest z ograniczeniami na import pewnych towarów (np. mięsa) z niektórych krajów. Wtedy dla takich krajów próg wejścia po stronie podaży jest nieskończenie wysoki.

Najwyższą formą rozwoju powszechnego rynku towarów i usług jest rynek integracji gospodarczej, np. wspólny rynek Unii Europejskiej, na którym państwa członkowskie mają wiele naturalnych przywilejów w porównaniu z nie-członkami UE. Historia Unii Europejskiej jest w istocie historią rozwoju wspólnego rynku. Rozpoczęła się ona od Wspólnoty Węgla i Stali, wspólnego rynku 6 państw (Belgii, Francji, Holandii, Luksemburga, RFN oraz Włoch) obejmującego tylko te dwa surowce strategiczne. Dziś, w 2011 roku, UE liczy 27 państw i jest najpotężniejszą organizacją gospodarczą świata. Tworzą one wspólny rynek około pół miliarda konsumentów (aktorów tego rynku po stronie popytu), na którym, zgodnie z jego ideą, mogą być swobodnie, bez żadnych ograniczeń, wymieniane cztery podstawowe dobra: kapitał (w naszej terminologii kapitał finansowy), praca, towary i usługi. Wprawdzie codzienna praktyka dobitnie pokazuje, że realizacja tej idei nie jest sprawą łatwą, ale dwa fakty wymagają szczególnego podkreślenia:

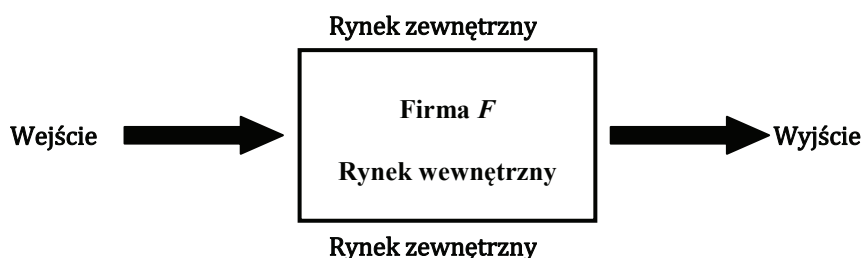
1. Jest wiele krajów aplikujących do UE, ale, jak dotąd, nikt z niej nie chce wystąpić.

2. Już od ponad 65 lat Europa, kontynent „od zawsze” targany wojnami i konfliktami narodowościowymi, żyje w pokoju i powiększa swój dobrobyt. W powszechnej opinii wkład wspólnego rynku UE, tej sieci formalnych i nieformalnych relacji między poszczególnymi państwami, firmami, ludźmi itp. w utrzymanie pokoju w Europie jest bardzo istotny.

Dotychczas mówiliśmy o powstawaniu nowych rynków w miarę rozwoju gospodarczo-społecznego. Jest oczywiste, że znikanie lub obumieranie rynków jest negacją lub odwrotną stroną procesu powstawania nowych rynków. Na początku ubiegłego wieku w Stanach Zjednoczonych istniał ogromny rynek brył lodowych, dostarczanych do domów przez bardzo rozbudowaną sieć dostaw. Wynalazek lodówki i jej upowszechnienie zlikwidowało ten rynek. Wielu Czytelników jest naoczniymi świadkami pojawienia się i zniknięcia rynku magnetowidów i magnetofonów taśmowych (szpulowych).

2.2. Rynek wewnętrzny

Rynek omówiony w poprzednim podpunkcie znajduje się na zewnątrz szeroko rozumianej firmy F , którą zdefiniowaliśmy w [10]. Jest to zatem rynek zewnętrzny w stosunku do firmy F . Samą firmę jako taką możemy potraktować jako rynek wewnętrzny, jako rynek umowny, stworzony specjalnie dla analizy firmy F , wnętrza prostokąta na rys. 1.



Rys. 1. Firma F jako rynek wewnętrzny

Na tym rynku zarząd/kierownictwo firmy F przydziela (umownie kupuje) pracowników na odpowiednie miejsca pracy/stanowiska tak, aby odnieść sukces na rynku zewnętrznym, na którym ta firma działa. Rezultatem tego umownego kupna jest wynagrodzenie za pracę w firmie F danego pracownika, co jest wartością jego kapitału ludzkiego na rynku wewnętrznym, wartością tego kapitału tu teraz [10]. Inaczej rzecz ujmując, pracownik sprzedaje swoją pracę, ściśle rzecz biorąc swoje możliwości świadczenia pracy na danym stanowisku, tj. swoje kompetencje, doświadczenie, talent itp. i za to otrzymuje wynagrodzenie, które wynika z działania prawa popytu i podaży (1) na rynku wewnętrznym.

Rynek wewnętrzny różni się od rynku zewnętrznego tym, że zwykle działanie prawa (1) na nim jest dodatkowo warunkowane tradycją, regulaminami, pisanymi/niepisanymi zasadami itp., obowiązującymi tylko w tej firmie. Z faktu, że firma F działa na wolnym rynku, w gospodarce rynkowej wynika, że jej pracownicy uważają tu i teraz reguły zatrudnienia, tj. reguły rynku wewnętrznego.

nego, za uczciwe, sprawiedliwe i adekwatne. Oznacza to też, że każdy pracownik podpisuje umowę o pracę w firmie F jako wolny człowiek, bez żadnego przymusu, nacisku itp.

Zatem aktorami na rynku wewnętrznym po stronie podaży są pracownicy firmy F , a po stronie popytu – kierownictwo tej firmy. Co więcej, kierownictwo firmy F jest równocześnie organizatorem rynku wewnętrznego. Pracownicy oferujący swoją pracę na danym stanowisku tworzą wewnętrzny rynek pracy (dla danego stanowiska), a kierownictwo firmy F występuje jako potencjalny nabywca na tym rynku. Jeśli ten nabywca wybierze pracownika na dane stanowisko, to podpisuje z nim umowę o pracę (na tym stanowisku) i płaci za tę pracę wynagrodzenie. Jeśli kierownictwo firmy F nie znajdzie wśród pracowników tej firmy odpowiedniego kandydata na dane stanowisko, to zwykle rozpoczyna poszukiwanie takiego kandydata na zewnętrznym (w stosunku do firmy F) rynku pracy. Zatem w każdym z tych dwóch przypadków kandydat podpisując umowę o pracę zgadza się na to, że zasady rynku wewnętrznego są uczciwe, sprawiedliwe i odpowiednie, co jest istotą definicji 2.

3. Rynek emocji sportowych

Sport jest bardzo wymierną dziedziną życia i doskonałym przykładem dla nas, jak mierzyć lub szacować to, co na pierwszy rzut oka wydaje się absolutnie niemierzalne. Przykładowo, w skokach narciarskich spopularyzowanych przez Adama Małysza, zawodnicy są oceniani za długość skoku i styl. Początkujący kibic może mieć wątpliwości, że o ile można zmierzyć długość skoku, to czy można obiektywnie ocenić jego styl (jakość)? Okazuje się, że można, i co więcej oceny te ostatnio uwzględniają warunki, w jakich skok został wykonany (siła, kierunek wiatru itp.). Miliony kibiców oglądających skoki narciarskie uważa te oceny za sprawiedliwe, uczciwe i adekwatnie opisujące wysiłek i klasę skoczka, co jest istotą definicji 2. W tym punkcie na przykładzie koszykówki, siatkówki i piłki nożnej pokażemy jak działa rynek emocji (sportowych), który zdefiniujemy tak:

Definicja 3

Na rynku emocji (sportowych) jego organizator sprzedaje emocje kibicom.

Zatem organizator rynku emocji jest aktorem po stronie podaży, a kibice – aktorami po stronie popytu. Jak zobaczymy dalej na przykładzie tych trzech dyscyplin sportowych, organizator proponuje kibicom emocje (obejrzenie zawodów) za odpowiednią opłatę (bilet wstępu, opłata za korzystanie z telewizji itp.), wynikającą z działania prawa popytu i podaży (1) na rynku emocji, działania tu i teraz. Kibice mogą albo pasywnie oglądać zapowiedzi zawodów, np. na plakatach, w Internecie itp. (porównaj z window shopping w podpunkcie 2.1),

albo kupić bilety czy zapłacić abonament telewizji kablowej i być wtedy aktywnymi aktorami na rynku emocji po stronie popytu. I jeszcze uwaga: wszystkie opinie w tym punkcie są opiniami autora, który ogląda sport zwykle w telewizji.

3.1. Koszykówka

W powszechnej opinii NBA National Basketball Association jest przykładem dobrego organizatora rozgrywek ligi profesjonalnej koszykówki w USA i związanego z nimi rynku emocji. Między innymi NBA ustala zasady sprzedaży/zakupu profesjonalnych koszykarzy przez kluby. Z pewną przesadą można powiedzieć, że handel koszykarzami różni się od handlu niewolnikami tym, że ceny są o niebo wyższe, a koszykarzom podczas testów medycznych zagląda się nie tylko w zęby. NBA ustala limit pieniędzy, jakie klub może wydać rocznie na zakup nowych koszykarzy. W przypadku przekroczenia tego limitu klub płaci NBA bardzo wysoki podatek i pieniądze uzyskane tą drogą są wykorzystywane na szkolenie młodych koszykarzy. Limit wprowadzono, by uniknąć sytuacji, gdy bogaty klub kupuje najpierw najlepszych i najdroższych koszykarzy, a potem wygrywa (prawie) wszystkie mecze, co drastycznie obniża emocje z nimi związane i tym samym niszczy rynek emocji sportowych.

NBA wymaga, aby każda umowa między koszykarzem a jego klubem była jawna i transparentna. Koszykarzem NBA można zostać tylko po przejściu przez tak zwany draft – testy dla kandydatów na koszykarzy NBA. Po takich testach sporządza się ranking kandydatów, poczynając od najlepszego. Zakup najlepszych kandydatów jest proponowany w pierwszej kolejności najsłabszym klubom, które zajęły ostatnie miejsca w tabeli rozgrywek po zakończeniu ubiegłego sezonu. W postępowaniu tym chodzi o sportowe wzmocnienie słabszych klubów, aby mecze były jak najbardziej wyrównane, by dostarczały kibicom maksimum emocji.

NBA nie tylko reguluje zasady handlu koszykarzami, ale też modyfikuje/zmienia przepisy gry w koszykówkę tak, aby gra była ciekawa zarówno dla kibiców bezpośrednio oglądających zawody, jak też, a może przede wszystkim, dla widzów. Wszystkie zasady i regulaminy są spisane, jawne i ściśle egzekwowane. Relacje między sędzią a zawodnikiem mogą być tylko formalne i zgodne z przepisami. Koszykówka jest sportem kontaktowym, w którym zawodnicy walczą bezpośrednio ze sobą, na małej przestrzeni w ogromnym stresie, gdy emocje sięgają zenitu, co kibice bardzo lubią. Rola sędziego sprowadza się do pilnowania, aby była to nawet ostra walka, ale zawsze walka sportowa w ramach przepisów. Na przykład krytykowanie przez zawodnika decyzji sędziego podczas meczu jest karane jako przewinienie techniczne (dwa dodatkowe rzuty wolne dla przeciwników i potem oni posiadają piłkę). Dwa przewinienia techniczne tego samego zawodnika w trakcie jednego meczu skutkują jego wykluczeniem (wyrzuceniem) z gry. Również po meczu zawodnik nie może publicznie (w telewizji, prasie itp.) krytykować decyzji sędziów, gdyż za

to wykroczenie grożą kary finansowe liczone w grubych tysiącach dolarów. W przepisach tego typu chodzi o autorytet sędziego jako funkcjonariusza NBA, którego decyzje są niezaskarżalne i nie podlegają apelacji. Oczywiście sędzia jest człowiekiem i może się pomylić, a sędziowie którym to się zdarza zbyt często, są eliminowani przez NBA, aby reguły gry na rynku emocji były zawsze uczciwe, sprawiedliwe i adekwatne, zgodnie z istotą definicji 2.

Koszykówka jest sportem, w którym nie ma remisów, co oznacza, że w każdym meczu lub jego w kolejnych dogrywkach musi być wyłoniony zwycięzca. Ponieważ poziom sportowy drużyn są bardzo wyrównane, to często dochodzi do niezwykle emocjonujących końcówek, gdy jeden rzut do kosza przesądza o zwycięstwie lub porażce. Przykładowo, dwie sekundy do końca regulaminowego czasu gry, to w koszykówce wieczność. Zwykle drużyna posiadająca piłkę bierze wtedy czas (przerwę), by ustalić ostatnią akcję. Chociaż na parkiecie w każdej z drużyn gra po pięciu koszykarzy i mecz to porównanie tu i teraz kapitałów społecznych tych drużyn [10] (współpracy między zawodnikami i ich współpracy z trenerem, wzajemnego zaufania, itp.), to rzut ostatniej szansy wykonuje jeden zawodnik, a jego wynik zależy od jego kapitału ludzkiego. Wynik tego rzutu zależy od wielu różnych rzeczy, takich jak profesjonalizm rzucającego, jego odporność na stres itp., a w końcu może zależeć też od łutu szczęścia, a szczęście, jak wiadomo, sprzyja lepszym. Historia koszykówki zna wiele takich przypadków, gdy równo z syreną kończącą zawody trafiano do kosza z ponad 20 metrów, lub w innych niewyobrażalnych sytuacjach.

Koszykówka to sport zespołowy i od współpracy zawodników w drużynie, od ich wzajemnego zrozumienia, zaufania itp. zależy sukces lub porażka w każdym meczu. Wiadomo, że meczów nie wygrywa „zlepek” nawet największych gwiazd koszykówki, z których „każda ciągnie w swoją stronę”. Zatem drużyną nie może być luźny zbiór pojedynczych graczy (kapitałów ludzkich, w naszej terminologii – patrz [10]). Gwiazdor musi chcieć i umieć, jak mówią sprawozdawcy „podzielić się piłką” (podać ją koledze, znajdującemu się na dobrej pozycji). Piękno koszykówki, jak zresztą każdego sportu zespołowego, polega na tym, że umiejętnie łączy on wysoki kapitał społeczny (współpraca, zaufanie, solidarność itp.) z wartościowym kapitałem ludzkim (talent, poświęcenie, odporność na stres itp.).

Podsumowując, NBA jest dobrym organizatorem rynku emocji (sportowych), bo udało mu się stworzyć reguły gry, które wszyscy aktorzy tego rynku (miliony widzów rocznie na trybunach i miliardy przed telewizorami na całym świecie oraz zawodnicy sędziowie, działacze itp.) uważają za uczciwe, sprawiedliwe i adekwatne, co jest istotą definicji 2. Zauważmy ogromne zróżnicowanie ekonomiczne aktorów na rynku emocji. Na jednym biegunie mamy koszykarzy NBA, którzy zarabiają bardzo dobrze, posiadają zwykle po kilka domów, kilkanaście samochodów i mają tzw. wino, kobiety i śpiew. Na drugim biegunie mamy kibiców koszykówki w biednym kraju np. w Polsce lub na Litwie, dla których miesięczny abonament telewizji kablowej jest istotną pozycją

w rodzinnym budżecie. Jest rzeczą wysoce zastanawiającą, co sprawia, że ci tak bogaci koszykarze wylewają wiadra potu na treningach i ryzykują swoje zdrowie (kontuzje) na meczach? Zdaniem autora, motorem postępowania koszykarzy nie są głównie pieniądze, chociaż „kasa” zawsze się liczy, ale potrzeba uznania, która jest naturalną potrzebą każdego człowieka [10].

NBA wynalazło reguły rynku emocji (patrz definicja 2), na którym uznanie milionów kibiców dla talentu, pracy, poświęcenia itp. danego koszykarza, uznanie wyrażane tu i teraz na każdym meczu, jest zdaniem tego koszykarza uczciwe, sprawiedliwe i adekwatnie go oceniające. Co więcej koszykarz „łaknie (go) jak kania dżdżu”. Gdyby było inaczej, to koszykarz zakończyłby karierę. Bezstronny obserwator musi przyznać, że na meczach NBA kibice oceniają nie rasę, religię, narodowość itp. poszczególnych koszykarzy, ale to, mówiąc ogólnie, jak pięknie oni grają. A jest co podziwiać, obserwując mężczyznę o wadze 120 kg i wzroście 210 cm poruszającego się z gracją i zwinnością baletnicy. Zdaniem autora, rynek emocji NBA jest miejscem, gdzie ekonomia i piękno podają sobie rękę.

Oczywiście rynek emocji NBA jest żywym tworem, ciągle modyfikowanym w miarę potrzeb, na którym trzeba codziennie rozwiązywać życiowe problemy. Obecnie (wiosna 2011) na rynku NBA narasta konflikt o zarobki koszykarzy. W trosce o przetrwanie klubów najsłabszych ekonomicznie, NBA ogranicza procentowo część przychodów klubu, które można wypłacać koszykarzom. Koszykarze, oczywiście, chcą ten procent (udział) zwiększyć (to jeszcze jeden dowód na to, że „kasa” się liczy). Podobne spory w przeszłości prowadziły nieraz do kilkumiesięcznego lokautu NBA (odłożenia początku rozgrywek w danym sezonie o kilka miesięcy), ale w końcu zawsze kończyły się kompromisem.

3.2. Siatkówka

Wybraliśmy tę dyscyplinę, aby omówić przypadek, gdy nie udało się stworzyć reguł rynku emocji, które aktorzy tego rynku zgodnie z istotą definicji 2 uważaliby za uczciwe, sprawiedliwe i adekwatne.

Mistrzostwa Świata w Siatkówce Mężczyzn w 2010 roku we Włoszech zostały zorganizowane przez włoską i światową federację tego sportu jako organizatorów rynku emocji na tym turnieju. Podział uczestników na grupy i zasady awansu do dalszych rozgrywek zostały tak zaprojektowane, że stwarzały łatwiejszą drogę do finału Włochom i Brazylii. Co więcej, na początku Mistrzostw opłacało się przegrać mecz, bo potem droga do finału mogła być łatwiejsza. Zasady te powszechnie oceniono jako nieuczciwe i niesprawiedliwe, czemu kibice dawali wyraz bucząc podczas wielu meczów. Następne Mistrzostwa Świata będą rozegrane według nowych, miejmy nadzieję, sprawiedliwych zasad (reguł rynku emocji).

3.3. Piłka nożna

Angielska Premiership, liga najlepszych angielskich zawodowych klubów piłki nożnej, jest powszechnie uważana za jedną z najlepiej zorganizowanych lig sportowych na świecie zarówno w sensie emocji kibiców, jak też w sensie biznesowym. Jeszcze w latach 50. XX wieku, piłka nożna w Anglii była powszechnie traktowana jako sport klasy robotniczej, by nie powiedzieć motłochu. Do historii sportu lub raczej zezwierzęcenia kibiców piłkarskich przeszły wypadki na Stadionie Heysel w Brukseli w 1985 roku, kiedy to w trakcie meczu pomiędzy włoskim Juwentusem a angielskim Liverpooliem zginęło 39 osób. W odpowiedzi rząd Margaret Thatcher wydał bardzo ostre przepisy i wyeliminował chuligaństwo i bandytyzm ze stadionów piłkarskich w Wielkiej Brytanii. Po zaprowadzeniu porządku na stadionach, wyraźnie wzrosły wpływy klubów piłkarskich i to ze wszystkich źródeł (sprzedaż biletów i kart wstępu, reklamy w mediach i na stadionach, opłaty za prawa transmisji telewizyjnych, sprzedaż klubowych gadżetów, takich jak koszulki, emblematy itp.). Powstał w ten sposób nowy, ogromny rynek, który nazwaliśmy rynkiem emocji, gdyż to emocje kibiców są jego siłą napędową. O skali tego rynku może świadczyć fakt, że wpływy 20 największych klubów piłkarskich świata przekroczyły w 2005 roku 2 miliardy funtów brytyjskich i ciągle rosną.

Powyższy przykład wybraliśmy świadomie, by pokazać, że dobrze zorganizowany rynek emocji może rozwiązać wiele problemów społecznych i przyczynić się do rozwoju ekonomicznego kraju. W tym kontekście rozwiązywany przez dziesięciolecia problem chuligaństwa i bandytyzmu na polskich stadionach, jak też problem korupcji (kupowania meczów) w polskim piłkarstwie, to zdaniem autora, dwie strony tej samej monety. To przykład niemocy, by powiedzieć dosadniej, impotencji decyzyjnej, typowo polskiego, „jakoś to będzie” lub „chcieliśmy dobrze, a wyszło jak zawsze”. Pomimo zmiany ustroju, powszechnej demokratyzacji itp. brak jest w Polsce odważnych, którzy zbudowaliby rynek emocji piłkarskich w Polsce.

4. Rynek relacji

Jak wiemy, w naszym świecie są tylko relacje formalne i nieformalne i choć nie są one rozłączne, co praktycznie oznacza, że nie możemy oddzielić jednych relacji od drugich (patrz [9] oraz [10]), to wyjątkowo teraz ze względów metodologicznych, omówimy najpierw rynek relacji formalnych, a potem rynek relacji nieformalnych.

4.1. Rynek relacji formalnych

Relacje formalne wygodnie jest badać na dwóch różnych poziomach: w szeroko rozumianej firmie *F* lub na poziomie kraju (regionu), tj. w megafirmie MF (patrz [10]).

W typowej firmie przemysłowej lub usługowej relacje formalne (ich realizacja w praktyce) są regulowane zwykle wewnętrznymi regulaminami, tradycją, pisanymi oraz niepisanymi zasadami itp. W [10] omówiliśmy relację formalną między przełożonym (szefem) a podwładnym, nazwaną tam relacją podwładności lub krótko podwładnością. Jest to najważniejsza relacja formalna w typowej firmie przemysłowej/usługowej wynikająca ze stosunku pracy, z umowy o pracę zawartej między kierownictwem firmy a danym pracownikiem. Za realizację relacji formalnych pracownicy otrzymują wynagrodzenie (pensje, premie, nagrody itp.), zgodnie z regułami gry na rynku wewnętrznym danej firmy, co opisaliśmy w podpunkcie 2.2. Zatem możemy mówić o rynku relacji formalnych w danej firmie, którego organizatorem jest jej kierownictwo, ustalające wewnętrzne regulaminy (np. regulamin wynagradzania), zasady, przepisy itp., regulujące pracę/działanie tego rynku. To szeroko rozumiane kierownictwo firmy, a ściślej rzecz biorąc przełożeni, występują na rynku relacji formalnych po stronie popytu (chcą, aby dane relacje formalne były w praktyce analizowanej firmy realizowane) a pracownicy (podwładni) występują po stronie podaży, tj. są aktorami tego rynku po stronie podaży. Z faktu, że każdy pracownik, jako wolna i świadoma osoba podpisał umowę o pracę, tj. zgodził się na realizację w praktyce określonych relacji formalnych, wynika jego opinia, że reguły gry na tym konkretnym rynku relacji formalnych są uczciwe, sprawiedliwe i adekwatne, co jest istotą definicji 2.

W życiu publicznym (w megafirmie MF) rynek relacji formalnych jest regulowany aktami normatywnymi takimi jak konstytucja, kodeksy (np. kodeks ruchu drogowego, kodeks karny itd.), regulaminy, rozporządzenia itp. lub też przez wieloletnią tradycję, na przykład relacje formalne między proboszczem a jego parafianami. Z tych aktów normatywnych wynika między innymi obowiązek płacenia podatków, co można należy rozumieć jako interpretację prawa (1) na rynku relacji formalnych w danym kraju. Za pieniądze z podatków państwo (kierownictwo megafirmy MF) gwarantuje obywatelowi (podatnikowi) bezpieczeństwo, szkolnictwo, infrastrukturę (np. drogową) itp. Zatem, podobnie jak w przypadku firmy *F*, demokratycznie wybrana władza jest organizatorem rynku relacji formalnych i występuje na nim po stronie popytu, a obywatele po stronie podaży. Z aktu wyborczego wynika zgoda obywateli na konkretną władzę publiczną i ich opinia, że tu i teraz, ogólnie rzecz biorąc, reguły gry na rynku relacji formalnych w danej megafirmie MF są uczciwe, sprawiedliwe i odpowiednie.

4.2. Rynek relacji nieformalnych

Jako wprowadzenie do studiów nad relacjami nieformalnymi rozważmy dwie zaprzyjaźnione rodziny A oraz B. Możemy je potraktować jako pewien wyizolowany wycinek rzeczywistości, taki (mały) rynek relacji nieformalnych. Wtedy te dwie rodziny są równocześnie organizatorami i jedynymi aktorami na tym rynku. Organizatorami w tym sensie, że to one ustalają reguły, zwykle niepisane, obowiązujące na tym rynku. Na przykład, jeśli rodzina A zaprosi B na obiad do wystawnej restauracji, to rodzina A w tej chwili działa po stronie podaży (oferuje obiad w restauracji), a rodzina B po stronie popytu (może to zaproszenie przyjąć lub też jakoś się wymówić). Zwykle rodzina A oczekuje rewanżu, że po pewnym czasie B zaprosi A na podobny obiad i wtedy rodziny zamienią się rolami: rodzina B wystąpi po stronie podaży, natomiast A – po stronie popytu. Gdyby w rewanżu rodzina B zaprosiła A do baru mlecznego, to rodzina A nie może podać B do sądu i zwykle wtedy taka przyjaźń się kończy a ten rynek relacji nieformalnych znika, mimo gorących zapewnień B, że jest to ostatni bar mleczny w III/IV RP (niepotrzebne skreślić).

Ta historyjka pozwala nam zdefiniować rynek emocji nieformalnych.

Definicja 4

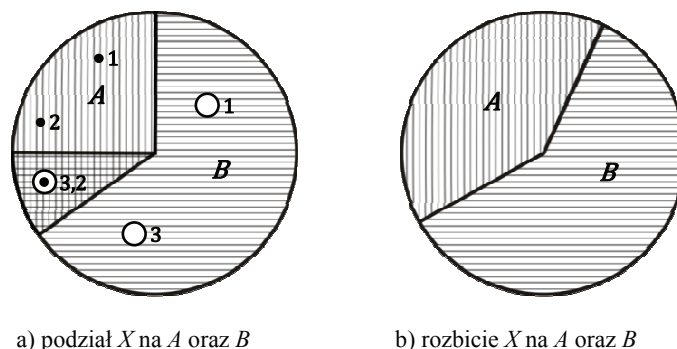
Wszyscy aktorzy rynku relacji nieformalnych definiują lub modyfikują jego reguły, są zatem jego organizatorami, jak również wymieniają się rolami, tj. występują raz po stronie podaży, a raz po stronie popytu.

Oczywiście, bezzasadne jest wymaganie wymiany na przemian ról po stronie popytu i po stronie podaży. Podobnie altruizm (np. rodzina A zaprasza zawsze B na obiady) jest wyjątkiem, potwierdzającym regułę. Już z definicji 4 Czytelnik może wywnioskować, dlaczego badanie rynku relacji nieformalnych jest tak trudne.

5. Istota zasady ortogonalności

By przybliżyć istotę zasady ortogonalności, rozważmy wszystkie, ale to absolutnie wszystkie wartości danej firmy F w chwili t . W [9] pokazaliśmy, że wygodnie jest czas t podzielić na trzy okresy: „dziś” podejmujemy decyzje na podstawie informacji z „wczoraj”, po to, aby odnieść sukces „jutro”. Inaczej mówiąc, analizujemy wycinek rzeczywistości X (wszystkie wartości firmy F) w chwili t . Fakt, że analizujemy wszystkie wartości, będziemy na rysunkach oznaczać za pomocą koła. Założmy, że analiza wszystkich wartości tego wycinka rzeczywistości jest dla nas zbyt skomplikowana i dlatego decydujemy się podzielić X na dwie formy (kategorie, pojęcia, wymiary itp.) A oraz B . Wtedy

możliwe są tylko dwa przypadki przedstawione na rys. 2: albo forma A , zakreskowana pionowo, zachodzi na formę B , zakreskowaną poziomo, co przedstawiono na rys. 2a jako część koła zakreskowaną w kratkę, albo jak na rys. 2b, formy te są rozłączne, tworząc rozbicie X na A oraz B . Różnicę między podziałem a rozbiciem można zobrazować następującym przykładem: jeżeli talerz upuszczony na podłogę rozbije się, to mamy jego rozbięcie na co najmniej dwie części, a nie podział. Inaczej mówiąc, każde rozbięcie jest podziałem, ale nie każdy podział jest rozbięciem. Nieraz rozbięcie jest nazywane podziałem rozłącznym.



Rys. 2. Różnica między podziałem a rozbięciem

Założmy dalej, że wyznaczamy wartość najpierw formy A , dodając wartości jej poszczególnych elementów (pozycji bilansowych) oznaczonych czarnymi kółkami, a później formy B , dodając wartości elementów oznaczonych białymi kółkami. Na rys. 2a drugi element formy B pokrywa się, jest tożsamy, z trzecim elementem A , a więc mamy tutaj przykład naruszenia podstawowej zasady rachunkowości, która mówi, że każdy element z analizowanego wycinka rzeczywistości jest uwzględniany w bilansie (analizie, wycenie itp.) jeden i tylko jeden raz. Oczywiście, taki przypadek nie może zaistnieć na rys. 2b, gdy formy są rozłączne. W naukach społecznych, takich jak: ekonomia, zarządzanie, rachunkowość, socjologia czy nauki polityczne, nie potrafimy dziś przedstawić analizowanego wycinka rzeczywistości X jako rys. 2a albo 2b, a więc trzeba to postępowanie opisać werbalnie (słowami), na przykład w formie zasady.

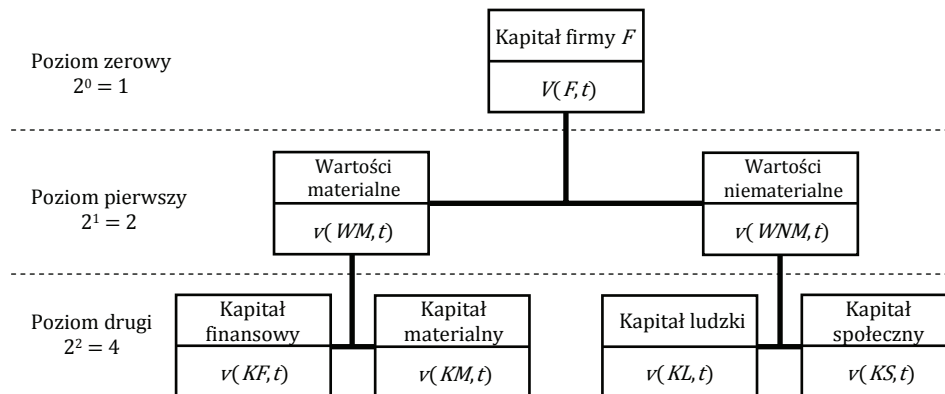
Definicja 5. Zasada ortogonalności

Dwa pojęcia (formy, kategorie, wymiary, koncepcje, procesy itp.) A oraz B są ortogonalne lub rozłączne wtedy i tylko wtedy, kiedy istnieje obiektywna, prosta, jednowymiarowa reguła decyzyjna typu tak-nie, pozwalająca zdecydować o przynależności dowolnego elementu/obiektu z rozważanego wycinka rzeczywistości X albo do A , albo do B .

Omówimy teraz cztery warunki, jakie nałożyliśmy na regułę decyzyjną w zasadzie ortogonalności.

Wymaganie, aby reguła decyzyjna była obiektywna oznacza, że wynik jej stosowania nie zależy ani od osoby stosującej tę regułę, ani też od czasu i miejsca, gdzie to zachodzi. W naszym rozumieniu reguła decyzyjna jest prosta, jeżeli można ją zdefiniować w jednym lub dwu zdaniach w potocznie używanym języku, bez odwoływania się do specjalistycznej terminologii. Na przykład, niech X będzie bilansem danej firmy sporządzanym przez księgowych. Wtedy nie jest prostą reguła decyzyjna opisana na, powiedzmy, dziesięciu stronach, której zrozumienie wymaga studiowania podręczników akademickich. Podobne założenia są podstawą wymagań, aby reguła decyzyjna była jednowymiarowa i typu tak-nie. Kontynuując przykład z księgowymi, zdaniem autora, dzisiaj i w dającej się przewidzieć przyszłości, trudno jest wymagać od nich stosowania analizy wielokryterialnej (wielowymiarowej) czy też rachunku zbiorów rozmytych.

W [9] (patrz również rozdział 2 oraz 3 w [10]) stosując zasadę ortogonalności pokazaliśmy, że wartość danej firmy F można zawsze rozbić na wartości materialne i niematerialne. Co więcej, te pierwsze można rozbić na kapitał finansowy i kapitał materialny (rzeczowy) rozważanej firmy, natomiast wartości niematerialne firmy F można rozbić na jej kapitał ludzki i kapitał społeczny. Mamy zatem następującą strukturę (układ) wartości/kapitałów, przedstawioną na rys. 3.



Rys 3. Rozbicie kapitału (wartości) firmy F na cztery różne formy

Z naszych rozważań wynika, iż cztery wyżej zdefiniowane formy kapitału firmy F nie mają elementów wspólnych, tj. są rozłączne lub ortogonalne, możemy zatem dodawać wartości odpowiadających im kapitałów. Udowodniliśmy w ten sposób zasadniczy rezultat naszych badań, który sformułujemy jako następujący lemat:

Lemat 1

Dla dowolnej firmy F jej wartość $V(F, t)$ w danej chwili t jest równa sumie wartości jej czterech kapitałów, tj.

$$V(F, t) = v(KF, t) + v(KM, t) + v(KL, t) + v(KS, t) \quad (2)$$

dla każdego t z jej „wczoraj“, „dziś” oraz „jutro”.

Definicja 6

Zależność (2) nazywamy równaniem fundamentalnym i jest to równość typu tu i teraz.

Równanie fundamentalne mówi, że w gospodarce rynkowej, w stanie równowagi, gdy popyt równoważy podaż, gdy obowiązuje prawo (1), wartość firmy w dowolnej chwili t z jej „wczoraj”, „dziś” oraz „jutro”, jest równa sumie wartości czterech i tylko czterech form kapitału wyznaczonych dla tej chwili, a mianowicie kapitału finansowego, materialnego, ludzkiego oraz społecznego. Zależność (2) nazwaliśmy równaniem fundamentalnym, gdyż jest ono podstawą teoretyczną, fundamentem naszych badań nad kapitałem ludzkim i kapitałem społecznym (patrz [10] oraz [4-9]). Badania te są intensywnie prowadzone w ostatnich 10-15 latach (patrz [1-3] oraz [11]).

Jak wiemy [9], kapitał społeczny tworzą formalne i nieformalne relacje między pracownikami firmy F . Relacja jest funkcją dwóch argumentów/zmiennych, a więc, aby zaistniała taka relacja, firma F musi zatrudniać co najmniej dwie osoby. Możemy więc sformułować następujący lemat:

Lemat 2

W jednoosobowej firmie nie ma kapitału społecznego, tj. $v(KS, t) = 0$ dla każdego t .

Inaczej rzecz ujmując, kapitał społeczny jest tam, gdzie są relacje formalne i/lub nieformalne między co najmniej dwiema (różnymi) osobami, nośnikami kapitału ludzkiego [10]. Ponieważ w firmie jednoosobowej, z definicji, pracuje/jest zatrudniona tylko jedna osoba, jej właściciel, to w tej firmie (wewnątrz tej firmy) nie może istnieć kapitał społeczny. Oznacza to, że wartość kapitału społecznego tej firmy równa się zero.

Wynik ten został opublikowany po raz pierwszy w marcu 2006 roku [4] i wzbudził spore kontrowersje. Krytycy tego rezultatu uważali, że skoro sukces jednoosobowej firmy istotnie zależy od relacji jej właściciela z klientami, włączając w to jego/jej relacje towarzyskie z klientami, to taka firma ma niezerowy kapitał społeczny. Błąd tego rozumowania polega na braku precyzji w definicji kapitału społecznego. Jeżeli mówimy o kapitale społecznym firmy F , to formalne i/lub nieformalne relacje, z których on się składa, muszą być

wewnątrz tej firmy. Fakt, na ile właściciel tej firmy jest, mówiąc po angielsku, social ze swoimi klientami, uwzględniamy w wycenie/szacowaniu jego/jej kapitału ludzkiego, a więc tu nic nie ginie.

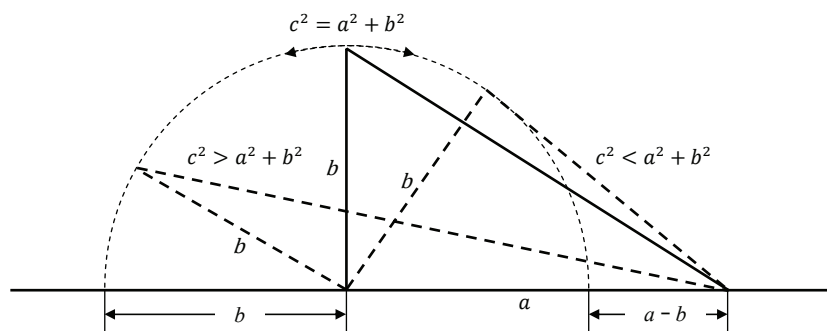
Należy z całą mocą podkreślić, że nie jest to zwykły zabieg techniczny przetrzucenia danej pozycji bilansowej z jednego rozdziału bilansu (kapitał społeczny) do drugiego (kapitał ludzki). W analizie jednoosobowej firmy F ta zaleta właściciela jest cechą pojedynczej istoty ludzkiej, właśnie tego nośnika kapitału ludzkiego, a więc jest to element (składowa) tego i tylko tego kapitału.

6. Twierdzenie Pitagorasa w naukach społecznych

W tym punkcie pokażemy, że z lematu 2 wynikają twierdzenia podobne do twierdzenia Pitagorasa, doskonale znanego ze szkoły średniej. Jak wiadomo, trójkąt można zadać/określić na wiele różnych sposobów, np. podając boki a oraz b i kąt α między nimi zawarty. Wtedy bok c tego trójkąta jest jednoznacznie określony, tj. jego długość można zawsze obliczyć. Istnieje wiele równoważnych sformułowań twierdzenia Pitagorasa, z których my wybierzemy następujące sformułowanie. Jego graficzną interpretację przedstawiono na rys. 4.

Twierdzenie Pitagorasa

1. Dla $\alpha = 90^\circ$ istnieje jeden i tylko jeden trójkąt prostokątny i wtedy $a^2 + b^2 = c^2$.
2. Dla $\alpha > 90^\circ$ istnieje nieskończenie wiele trójkątów rozwartokątnych i wtedy $a^2 + b^2 < c^2$.
3. Dla $\alpha < 90^\circ$ istnieje nieskończenie wiele trójkątów ostrokątnych i wtedy $a^2 + b^2 > c^2$.



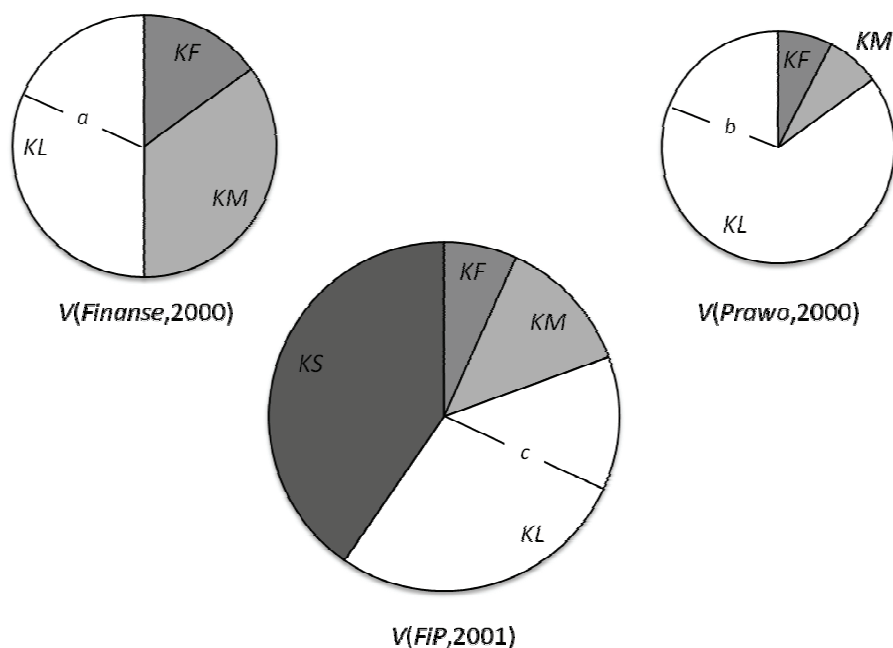
Rys. 4. Graficzna interpretacja twierdzenia Pitagorasa

Każde z twierdzeń analogicznych do twierdzenia Pitagorasa zostanie wprowadzone za pomocą przykładu, opisującego badany wycinek rzeczywistości, gdzie tu i teraz popyt równoważy podaż (1), skąd wynika równanie fundamentalne (2) i lemat 2.

6.1. Ekonomia i zarządzanie

Dwaj właściciele jednoosobowych firm Finanse oraz Prawo postanowili na początku 2001 roku połączyć swoje wysiłki i utworzyć dwuosobową firmę Finanse i Prawo, w skrócie FiP. Na rys. 5 przedstawiono wartości tych jednoosobowych firm przed ich połączeniem, pod koniec 2000 roku. Pola kół o promieniach odpowiednio a oraz b są proporcjonalne do wartości tych firm. Niech K będzie współczynnikiem proporcjonalności.

Zgodnie z lematem 2, firmy te nie miały wtedy kapitału społecznego, a więc na ich wartości składały się kapitał finansowy, materialny i ludzki. Z tego rysunku wynika, że firma Prawo, tj. jej właściciel, była biedna jak mysz kościelna, gdyż posiadała znikomy kapitał finansowy i materialny i prawie całą jej wartość stanowił kapitał ludzki właściciela (jego wiedza prawnicza, doświadczenie, urok osobisty itd.).



Rys. 5. Relacje między kapitałami łączących się firm

Z chwilą połączenia firm, na początku 2001 roku, w dwuosobowej firmie FiP pojawił się kapitał społeczny, tj. $v(KS, FiP, 2001) = d > 0$. Ponieważ to połączenie było ekonomicznie uzasadnione, to wystąpił efekt synergii, a mianowicie, wartość FiP była istotnie (ostro) większa niż suma prosta wartości tworzących ją firm

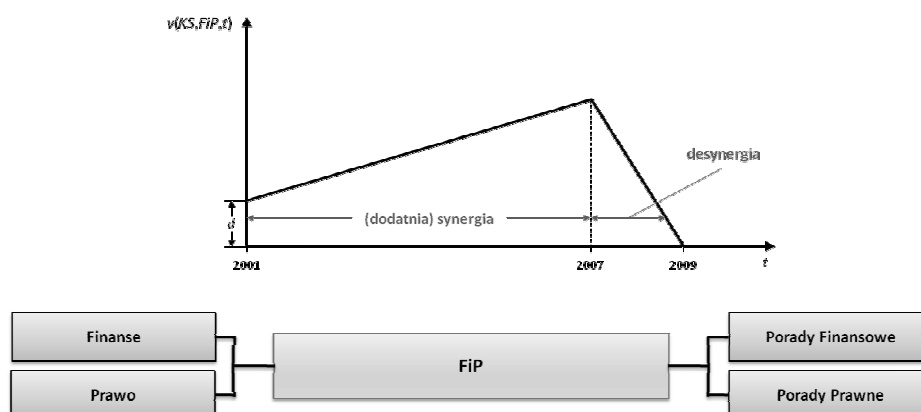
$$V(FiP, 2001) > V(Finanse, 2000) + V(Prawo, 2000)$$

Zauważmy, że z uwagi na wspomnianą synergię pole koła odpowiadającego firmie FiP jest większe niż suma pól kół dwóch tworzących ją firm, czyli

$$K\pi c^2 > K\pi a^2 + K\pi b^2$$

Początkowo firma FiP świetnie się rozwijała, rosła jej wartość oraz wartość jej kapitału społecznego, co przedstawiono na rys. 6, ale w 2007 roku wspólnicy zaczęli się kłócić i to coraz ostrzej, a jej kapitał społeczny, tj. jego wartość, zaczął maleć. Mamy wtedy do czynienia ze zjawiskiem de-synergii, jako zaprzeczenie (negację) opisanego wyżej zjawiska synergii. W końcu, w 2009 roku firma FiP zbankrutowała, rozpadła się na dwie jednoosobowe firmy Porady Finansowe oraz Porady Prawne. Wtedy, jak wynika z lematu 2, wartość kapitału społecznego tej firmy była równa zeru, tj. $v(KS, FiP, 2009) = 0$. Gdy wspólnicy kłócili się, to do zera zmalał nie tylko kapitał społeczny, ale również pogarszała się sytuacja finansowa firmy FiP. Pomimo sprzedaży wszystkiego, co składało się na kapitał materialny firmy (meble, komputery, itp.), co można było sprzedać, wpadła ona w dług, zatem jej kapitał finansowy stał się ujemny, tj. $v(KF, FiP, 2009) < 0$. Prócz tego FiP musiała zapłacić czynsz za wynajmowane pomieszczenia biurowe o jeden miesiąc dłużej niż okres wynajmu, czyli jej kapitał materialny stał się również ujemny, tj. $v(KM, FiP, 2009) < 0$.

Ze względów metodologicznych przyjęliśmy, że kapitał społeczny – $v(KS, FiP, t)$ na rys. 6 rósł i malał liniowo. Rysunek ten ilustruje dobrze znaną życiową prawdę, że zgoda buduje (powoli), a niezgoda rujnuje (szybko).



Rys. 6. Zmiana wartości kapitału społecznego firmy FiP w latach 2001-2009

Możemy teraz sformułować twierdzenie Pitagorasa w ekonomii i zarządzaniu:

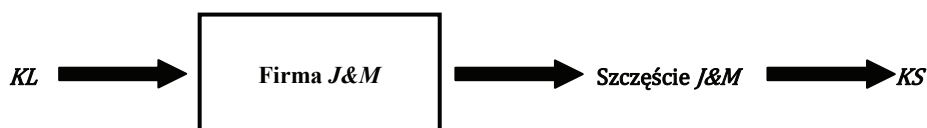
1. Istnieje tylko jedna i tylko jedna forma niezależności ekonomicznej lub ortogonalności dwóch szeroko rozumianych firm.
2. Istnieje nieskończenie wiele form synergii (współpraca ekonomiczna, kooperacja biznesowa itp.).
3. Istnieje nieskończenie wiele form desynergii (wojny ekonomiczne, spory kompetencyjne itp.).

Łatwo sprawdzić, że sformułowane wyżej twierdzenie zachodzi również w przypadku dwóch dowolnych państw, (krajów), tj. megafirm MF w naszej terminologii. Zatem istnieje analog twierdzenia Pitagorasa w naukach politycznych.

6.2. Socjologia i psychologia

Zasadniczym celem jest tu pokazanie na przykładzie, jak należy stosować lemat 2 w analizie kapitału ludzkiego i społecznego. W sposób całkowicie zamierzony wybraliśmy szczęście (małżeńskie) i miłość jako przedmiot tej analizy, by pokazać, że nasza teoria pozwala spojrzeć na powszechnie znane fakty z nieco innej perspektywy. To nowe spojrzenie świadczy o sile tej teorii, że pozwala ona analizować tak ulotne i nieokreślone zjawiska jak szczęście lub miłość. Analizę tę prowadzimy przy kilku upraszczających założeniach, które pozwalają lepiej przedstawić istotę rzeczy. Czynimy je świadomie, gdyż jest to opracowanie naukowe, a nie poradnik dla prowadzących kąciki złamanych serc w prasie kobiecej.

Rozważmy małżeństwo Jasia i Małgosi jako firmę dwuosobową (w skrócie J&M) przedstawioną na rys. 7, której celem istnienia/aktywności jest szczęście (małżeńskie). Z tego rysunku wynika, że firma J&M transformuje/przetwarza cztery formy kapitału na wejściu w szczęście na wyjściu. W tym punkcie będziemy analizowali zależność kapitału społecznego na wyjściu systemu (firmy) J&M od kapitału ludzkiego na jej wejściu. By uprościć tę analizę, ograniczymy kapitał społeczny (wyjście) do tylko jednej jego składowej, jaką jest w tym przypadku szczęście (małżeńskie), a wszystkie inne jego składowe świadomie pomijamy. Podobnie postępujemy z wejściem i ograniczamy kapitał ludzki do tylko jednej jego składowej – KL – miłości do osoby przeciwnej płci.



Rys. 7. Małżeństwo jako firma dwuosobowa J&M

Historia firmy J&M rozpoczyna się w dniu ślubu ($\hat{S}$) Jasia i Małgosi. Dla każdego t jej kapitał ludzki jest sumą ich kapitałów ludzkich

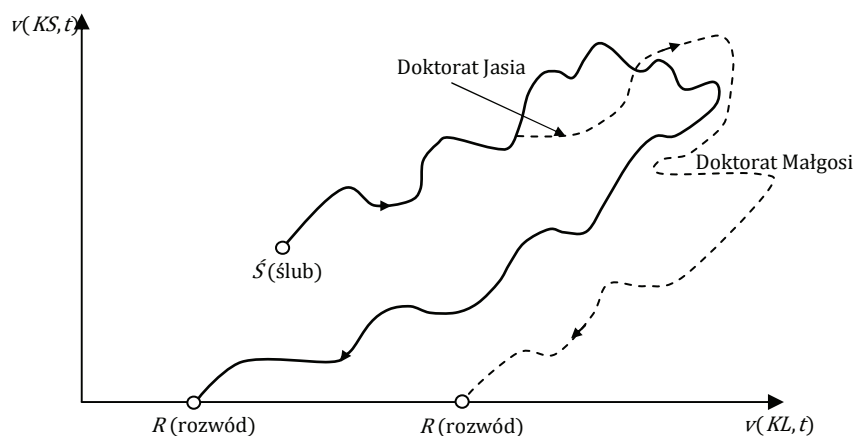
$$v(KL, J\&M, t) = v(KL, Jaś, t) + v(KL, Małgosia, t)$$

Możemy tak zrobić, ponieważ Jaś i Małgosia to dwie różne osoby i ich kapitały ludzkie są ortogonalne zgodnie z zasadą ortogonalności (definicja 5). Ponieważ młodzi znali się przed ślubem, to zarówno kapitał ludzki jak i kapitał społeczny firmy J&M w dniu ich ślubu był dodatni, tj.

$$v(KL, J\&M, t = \hat{S}) > 0 \text{ oraz } v(KS, J\&M, t = \hat{S}) > 0$$

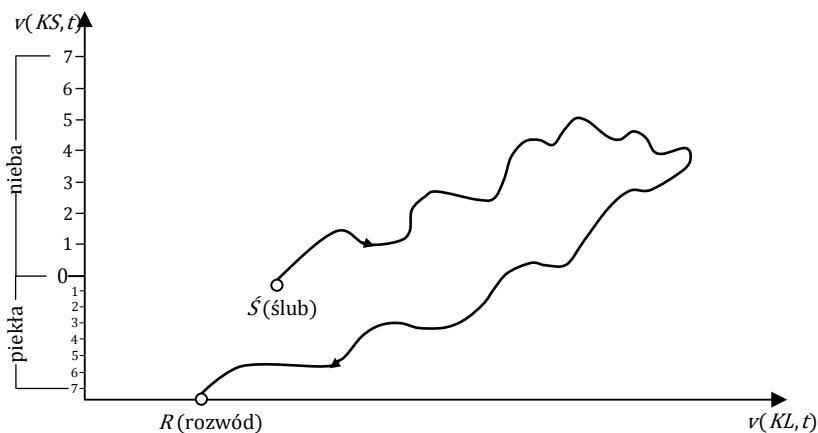
co widać na rys. 8, gdzie przedstawiono zależność (wartości) kapitału społecznego, tj. wyjścia systemu (firmy) J&M od (wartości) kapitału ludzkiego na wejściu tego systemu. Po ślubie młodzi byli szczęśliwi, rozmawiali o „górach i chmurach, o błocie i złocie”, a więc ich kapitały ludzki i społeczny, to znaczy ich wartości, rosły. Nikt nie wie dlaczego w pewnym momencie zaczęli się kłócić i to coraz ostrzej, z biciem talerzy, niszczeniem mebli itd. Z powodu tych kłótni ich kapitał ludzki i społeczny zaczął maleć, podobnie jak ich kapitał materialny i finansowy, bo wizyty u lekarzy, psychologów, prawników itp. dużo kosztują.

W dniu R nasi bohaterowie się rozwiedli, firma J&M przestała istnieć i rozpadła się na dwie jednoosobowe firmy „Jaś z odzysku” i „Małgosia z odzysku”. Zgodnie z lematem 2 kapitał społeczny J&M w dniu R miał wartość zero, ale już jej kapitał ludzki, jak też kapitały ludzkie dwóch nowo powstałych firm jednoosobowych były w tym dniu dodatnie (patrz rys. 8) Potwierdzeniem tego jest fakt, że ludzie rozwiedzeni wstępują w kolejne związki małżeńskie. Linią przerywaną pokazano na tym rysunku jak doktoraty Jasia i Małgosi mogły wpływać na ich szczęście małżeńskie.



Rys. 8. Zależność kapitału społecznego od kapitału ludzkiego w firmie J&M

Czytelnik zapewne zauważył, jak wiele możemy powiedzieć o wzajemnych zależnościach pomiędzy szczęściem małżeńskim (skutek, wyjście) a miłością (przyczyna, wejście), chociaż dziś nie wiem jak i w czym je mierzyć. Jeśli szczytem szczęścia jest „być w siódmym niebie”, a jego dnem jest „siódme nieszczęście (piekło)”, to można zaproponować, aby mierzyć szczęście tak, jak na rys. 9.



Rys. 9. Proponowane miary kapitału społecznego (szczęścia) w firmie J&M

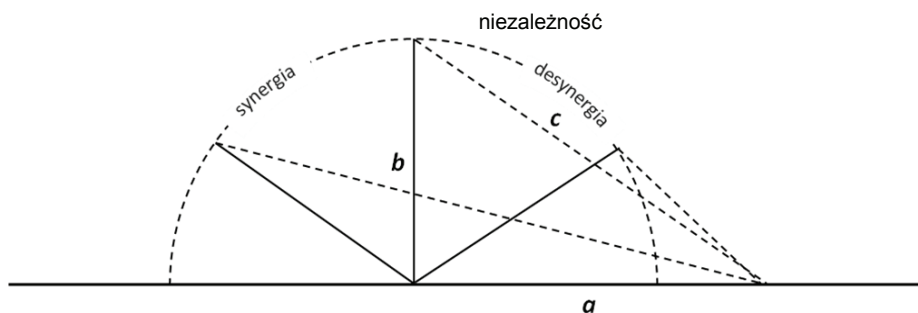
Z rysunku tego wynika, że, jak to zwykle w życiu bywa, Jaś i Małgosia mieli jeszcze daleko do „siódmego nieba”, a rozwód to coś znacznie gorszego niż „siódme nieszczęście”. Zgodnie z ludową mądrością, że „nieszczęścia

chodzą parami”, odległości między sąsiednimi „nieszczęściami” są znacznie mniejsze niż odległości między sąsiednimi „niebami”. Jest to też zgodne z powszechnym odczuciem, że „strata bardziej boli, niż zysk cieszy”.

Podobnie jak w poprzednim podpunkcie sformułujemy twierdzenie Pitagorasa w psychologii i socjologii:

1. Istnieje tylko jedna i tylko jedna forma niezależności lub ortogonalności dwóch samodzielnych istot ludzkich.
2. Istnieje nieskończenie wiele form synergii (miłość, przyjaźń, współpraca itp.).
3. Istnieje nieskończenie wiele form desynergii (konflikty, kłótnie, spory itp.).

Rysunek 10 jest graficzną interpretacją twierdzenia Pitagorasa w naukach społecznych. W podręczniku [10] pokazaliśmy bliski związek wielkiego twierdzenia Fermata z zasadą ortogonalności (definicja 5). Dwa powyższe twierdzenia również pokazują ten związek i tym samym mogą być traktowane jako jeszcze jedna interpretacja graficzna tej zasady.



Rys. 10. Graficzna interpretacja Twierdzenia Pitagorasa w naukach społecznych

Podsumowanie

Ten artykuł można traktować jako początek nowego działu matematyki, który proponujemy nazwać matematyką relacji tu i teraz lub w skrócie matematyką titów. Jak pokazaliśmy, potraktowanie powszechnie znanej prawdy, że w gospodarce rynkowej popyt równa się podaży, jako relacji (równości) tu i teraz dało wiele ciekawych i oryginalnych wniosków. W powszechnej opinii prawdziwa matematyka rozpoczęła się od twierdzenia Pitagorasa, które ma bardzo głęboki sens filozoficzny i estetyczny. Twierdzenie to mówi, że to co w naszym życiu jest piękne (racjonalne, proste, ortogonalne, rozłączne itp.) ma elegancką postać matematyczną. Ktoś może powiedzieć, że Krzywa Wieża w Pizie jest też piękna, ale to jest wyjątek, który tu i teraz potwierdza poprzednie zdanie.

Matematyka tytułów ma swój odpowiednik twierdzenia Pitagorasa, podobnie jak teoria liczb, teoria prawdopodobieństwa czy statystyka. Autor jest gorącym zwolennikiem tezy, że nauka, jako poszukiwanie prawdy, jest jedna, a podział na dyscypliny, takie jak matematyka, ekonomia czy psychologia, acz dziś nieunikniony, poszedł w wielu przypadkach zbyt daleko, nieraz ocierając się o granice absurdu. W tym kontekście możemy spojrzeć na naukę jako poszukiwanie reguł bądź to reguł decyzyjnych, takich jak w zasadzie ortogonalności (definicja 5), które pozwalają nam lepiej opisać i zrozumieć świat albo reguł gry (rynkowej), które pozwalają nam rozsądniej ten świat tu i teraz zorganizować (patrz definicja 2).

Literatura

1. Edvinsson L. and Malone M.S. (1997). *Intellectual Capital. The Proven Way to Establish your Company's Real Value by Measuring its Hidden Brainpower*. HaperBusiess, London.
2. Li P.P. (2007). Social Tie, Social Capital, and Social Behavior: Toward an Integrative Model of Informal Exchange. *Asia Pacific J. Management*, 24, s. 227-246.
3. Pawar M. (2006). Socialcapital? *The Social Science Journal*, 43, s. 211-226.
4. Walukiewicz S. (2006). *Systems Analysis of Social Capital at the firm Level*. Working Paper WP-1-2006, Systems Research Institute, Warsaw.
5. Walukiewicz S. (2006). Trzy modele do analizy kapitału społecznego. W: BOS 2006 – Wiedza systemowa. Red. J. Stachowicz, A. Straszak, S. Walukiewicz. IBS PAN, Warszawa, s. 25-40.
6. Walukiewicz S. (2007). Four Forms of Capita and Proximity. Working Paper WP-3-2007, Systems Research Institute, Warsaw.
7. Walukiewicz S. (2008). Piękno liczby cztery (w naukach społecznych). Working Paper WP-2-2008, Instytut Badań Systemowych PAN, Warszawa.
8. Walukiewicz S. (2008). The Dimensionality of Capital and Proximity. *Proceedings of ERSa 2008*. Liverpool, August 27-30.
9. Walukiewicz S. (2010). Zasada ortogonalności I przykłady jej zastosowań. W: *Modelowanie preferencji a ryzyko '09*. Red. T. Trzaskalik. AE, Katowice, s. 279-302.
10. Walukiewicz S. (2010). *Kapitał ludzki*. Skrypt akademicki. Instytut Badań Systemowych PAN, Warszawa.
11. Young G.K. (2007). *A Survey on Intangible Capital*. CEI Working Paper Series, No. 2007-10, Hitotsubashi University, Tokyo, Japan.

MATHEMATICS OF HERE AND NOW RELATIONS***Summary**

Everyone knows that in the market economy under equilibrium conditions:

$$\text{Demand} = \text{Supply} \quad (1)$$

This equality is an example of a here and now relations as it does not hold true, for instance, in North Korea and Cuba nowadays and it did not in Europe during the Second World War. So the same sign (=) has in (1) much narrower sense than in for $2+3=5$ which holds true always and everywhere.

We will study here (1) in detail as a base for mathematics of here and now relations. In fact, (1) says that the (aggregated) value of demand equals the value of supply here and now. Using here and now relations we provide a very general definitions of “value” and “market” notions. From (1) we derive the Fundamental Equation [4] saying that the value of a broadly defined company equals here and now the sum of four and only four its capitals: financial, physical, human and social. We also propose a very interesting theorem in analogy to the Theorem of Pythagoras.

* This study was sponsored through a grant UDA-POKL.04.02.00-00-083/08-00 “Human and Social Capital as new Academic Subjects”, awarded by the Ministry of Science and Higher Education of Poland.

Aneta Anna Wiktorzak

Instytut Badań Systemowych PAN w Warszawie

Państwowa Wyższa Szkoła Informatyki i Przedsiębiorczości w Łomży

METODY POMIARU OSIĄGNIĘĆ NA KOLEJNYCH SZCZEBŁACH EDUKACJI

Wprowadzenie

Proces kształtowania kapitału ludzkiego zaczyna się już w momencie narodzin i trwa praktycznie przez całe życie. Nie ulega jednak wątpliwości, że edukacja szkolna w największym stopniu wpływa na jego rozwój. To właśnie szkoła już od najmłodszych lat powinna zachęcać do powiększania wiedzy. Edukacja szkolna rozumiana jest przez nas jako złożony proces, który rozpoczyna się w szkole podstawowej, trwa w gimnazjum i szkole średniej, a swój koniec ma na uczelni.

Celem opracowania jest zwrócenie szczególnej uwagi na wkład szkół i uczelni w rozwój kapitału ludzkiego danej osoby oraz opis metod pomiaru wartości kapitału ludzkiego. Metody te wykorzystują wyniki państwowych egzaminów zewnętrznych w szkołach oraz wyniki na studiach liczone w punktach ECTS (European Credit Transfer System) – Europejskiego Systemu Transferu Punktów. W opracowaniu podajemy również krótką charakterystykę i opis działania sztucznej sieci neuronowej, którą wykorzystano w obliczeniach i analizie danych. Omówiono w skrócie regresję liniową, jako jedną z metod analizy współzależności zjawisk masowych. Na koniec przedstawimy analizę wyników 156 osobowej grupy studentów/uczniów, na przykładzie której pokażemy, jak ta teoria powinna być stosowana w praktyce.

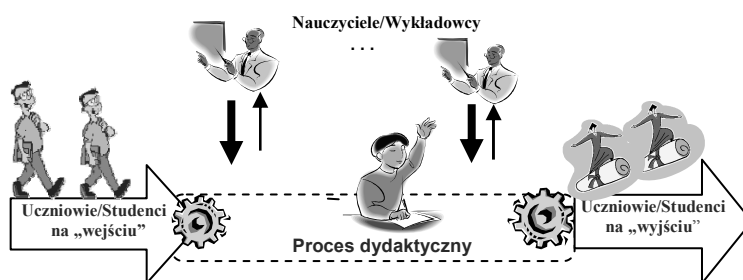
1. Proces kształtowania kapitału ludzkiego

Pojęcie kapitał ludzki (human capital) pojawiło się po raz pierwszy jako przedmiot naukowych dyskusji w 1975 roku za sprawą noblisty z zakresu ekonomii Gary’ego S. Beckera [1]. W polskim piśmiennictwie funkcjonowało ono najpierw jako społeczny koszt pracy żywej, a użyte zostało po raz pierwszy przez Domańskiego [2], który zdefiniował je jako „[...] zasób wiedzy,

umiejętności, zdrowia i energii witalnej, zawarty w danym społeczeństwie/narodzie”. W opracowaniu przyjmujemy definicję, że kapitał ludzki to zasób wiedzy, umiejętności, kompetencji, zdrowia i energii witalnej itp. zawarty w każdym człowieku traktowanym, jako oddzielna istota ludzka [13; 15]. Określa on między innymi zdolności tej istoty do pracy, zdolność do adaptacji do zmian zachodzących w jej otoczeniu oraz potrzebę podejmowania wyzwań i poszukiwania nowych rozwiązań. Tak rozwijany kapitał ludzki rozwijany jest przez całe życie.

Kapitał ludzki ucznia to wszystkie, ale to absolutnie wszystkie, zasoby niematerialne kojarzone tu i teraz z danym uczniem, jako samodzielną istotą ludzką [13]. W analizie kapitału ludzkiego danego ucznia wyróżniamy: szeroko rozumianą wiedzę, umiejętności, zdolności, talent, zdrowie, nastawienie do życia i otoczenia.

Do analizy i szacowania wartości kapitału ludzkiego ucznia/studenta wykorzystamy Wirtualną Taśmę Produkcyjną (WTP) zaproponowaną po raz pierwszy w 2006 roku przez Walukiewicza [9]. Na danej WTP uczniowie/studenci uczą się od nauczycieli/wykładowców poszczególnych przedmiotów w określonej kolejności, zakresie, obciążeniu dydaktycznym itp. Zatem na danej WTP realizuje się proces dydaktyczny, tak jak to pokazano na rys. 1.



Rys. 1. WTP dla procesu dydaktycznego

Należy z całą mocą podkreślić, że uczniowie/studenci są traktowani na WTP jako istoty ludzkie, które czują, przeżywają itp. WTP pozwala spojrzeć z nowej perspektywy na całą złożoność procesu edukacyjno-wychowawczego.

Niech $v(KL, S_i, t_k)$ oznacza wartość kapitału ludzkiego ucznia/studenta S_i w momencie zakończenia pewnego etapu nauki np. na koniec szkoły średniej, czy szkoły gimnazjalnej. Natomiast $v(KL, S_i, t_0)$ oznacza wartość kapitału ludzkiego ucznia/studenta S_i w momencie rozpoczęcia pewnego etapu nauki np. na początku szkoły średniej, czy w momencie rozpoczęcia nauki na studiach. Bezwzględny przyrost wartości kapitału ludzkiego ucznia/studenta S_i , oznaczony jako V_b , to różnica wartości tego kapitału na wyjściu i wejściu WTP realizującej jeden z etapów procesu edukacyjnego

$$V_b(KL) = v(KL, S_i, t_k) - v(KL, S_i, t_0)$$

Miara ta ma stosunkowo małą wartość poznawczą i w praktyce korzystamy ze względnych miar przyrostu wartości kapitału ludzkiego $V(KL)$, które mówią, jak dany uczeń/student (ściśle mówiąc przyrost/ubytek wartości jego kapitału ludzkiego) wygląda na tle klasy, grupy, szkoły, uczelni, regionu czy kraju. Miara ta zostanie rozwinięta w dalszej części opracowania.

Dziś mamy małe szanse na to, aby wartość kapitału ludzkiego ucznia/studenta była mierzona w jednostkach monetarnych, gdyż nie ma potrzeby i przyzwolenia społecznego na taką zmianę. Obecnie w Polsce jest ona mierzona niemal zawsze w stopniach, od jedynki do szóstki, z wyjątkiem szkół wyższych, gdzie obowiązuje 4-stopniowa skala ocen od dwójki do piątki i punkty ECTS. Niezależnie od tego w jakich jednostkach będziemy mierzyli wartość kapitału ludzkiego danego ucznia/studenta na wejściu i wyjściu WTP (w stopniach, punktach, procentach itp.), to można pokazać, że dla dostatecznie dużego n rozkład $v(KL, S_i, t)$ można aproksymować rozkładem normalnym lub rozkładem Gaussa, gdzie n jest liczebnością danej populacji.

2. Pomiar kapitału ludzkiego na kolejnych stopniach edukacji

Wyszkolenie stanowi istotny element kształtowania wartości kapitału ludzkiego. W związku z tym, nie powinien dziwić nas fakt, że przy próbie pomiaru wartości kapitału ludzkiego będziemy w głównej mierze zwracać uwagę na wyniki kształcenia. W naszych badaniach wykorzystamy wyniki egzaminów zewnętrznych. Dopiero teraz pojawiła się szansa kompleksowej analizy zmian poziomu pozyskiwanej wiedzy i umiejętności na poszczególnych szczeblach edukacji, a co za tym idzie zmian w podejściu do szacowania wartości kapitału ludzkiego poczynając od ucznia szkoły podstawowej po studenta szkoły wyższej. Egzaminy zewnętrzne, a właściwie sposób myślenia, z którego one wyrosły, stopniowo wywierały coraz większy wpływ na szkoły, instytucje edukacyjne, nauczycieli, uczniów, rodziców. W artykule stawimy nawet tezę, że znaczenie tego sposobu myślenia i jego wpływ na system edukacji rośnie wraz ze świadomością znaczenia budowania kapitału ludzkiego ucznia, studenta czy obywatela. System egzaminów zewnętrznych spowodował tak głębokie zmiany, że w praktyce mamy do czynienia z kształtowaniem się nowego, jakościowo innego podejścia do edukacji, a tym samym do badań nad możliwościami pomiaru kapitału ludzkiego.

2.1. Interpretacja wyników egzaminów

Najpierw opiszemy system egzaminacyjny w Polsce, tj. jego kolejne szczeble. W punkcie tym również wyjaśniamy sposób przeliczania osiągnięć na studiach w systemie ECTS, obowiązującym na uczelniach wyższych. Tabela 1 przedstawia najważniejsze cechy systemu egzaminowania zewnętrznego na kolejnych stopniach edukacji po: szkole podstawowej, szkole gimnazjalnej i szkole średniej oraz sposób oceniania studentów z wykorzystaniem punktów ECTS.

Tabela 1

System egzaminów zewnętrznych w Polsce

Szkoła	Egzamin	Od roku	Forma	Punkty
Podstawowa	Szóstoklasisty	2002	Jeden egzamin pisemny	Max 40
Gimnazjum	Egzamin gimnazjalny	2002	Trzy egzaminy pisemne*:	Max 150
Ponad-gimnazjalna	Matura	2005	Trzy przedmioty obowiązkowe**:	Max 170
Studia	Egzaminy i zaliczenia		Egzaminy, zaliczenia, laboratoria itp.	Max 300 pkt. 60 pkt. ECTS na rok***

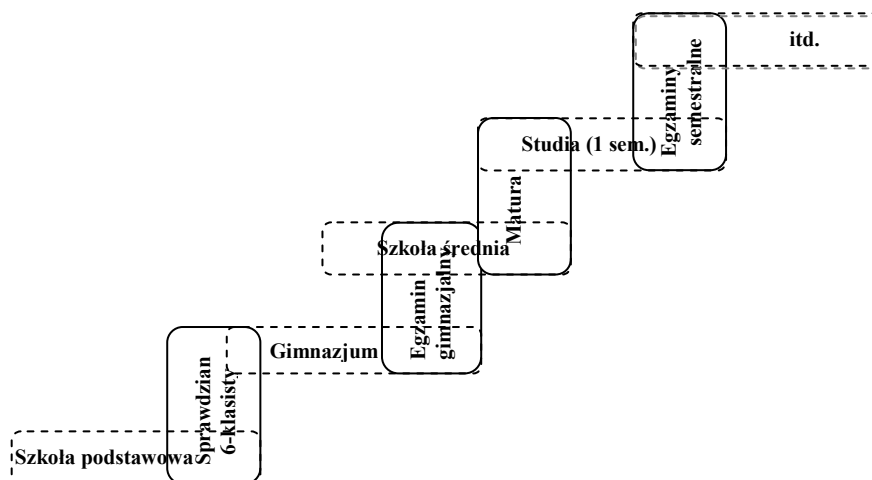
* Od 2009 roku na egzaminie gimnazjalnym uczniowie zdają trzecią część – językową, również za 50 pkt.

** W naszych badaniach brane są pod uwagę tylko przedmioty obowiązkowe, wyniki części pisemnej, na poziomie podstawowym.

*** Wszystkie egzaminy w sesji zdane na ocenę bardzo dobrą razy 30 pkt. ECTS dają maksymalnie 150 pkt. w semestrze, czyli 300 pkt. w roku akademickim.

Należy podkreślić, że wyniki uzyskane przez ucznia/studenta na wyjściu (na koniec danego szczebla edukacji) są jednocześnie danymi wejściowymi na kolejnym stopniu kształcenia. Na przykład wynik uzyskany na egzaminie szóstoklasisty są danymi wejściowymi do procesu kształcenia w gimnazjum, jak również są danymi wejściowymi w analizie wyników egzaminu gimnazjalnego (dane wyjściowe). Kolejno wyniki z egzaminu gimnazjalnego są danymi wejściowymi w procesie kształcenia w szkole średniej i są podstawą do analizy wyników maturalnych. Wyniki jakie uczeń uzyskał na maturze są danymi wejściowymi do analizy wyników uzyskiwanych przez studentów na studiach.

Idąc za myślą Walukiewicza [13] możemy sobie wyobrazić taką uogólnioną wirtualną taśmę produkcyjną, gdzie „produkty” z jednej taśmy przechodzą na kolejną. Przejście przez wszystkie fragmenty tej taśmy kojarzymy z przejściem ucznia/studenta przez kolejne poziomy edukacji od szkoły podstawowej, przez gimnazjum, szkołę średnią do studiów (rys. 2).



Rys. 2. Wirtualna taśma produkcyjna w całym procesie edukacyjnym

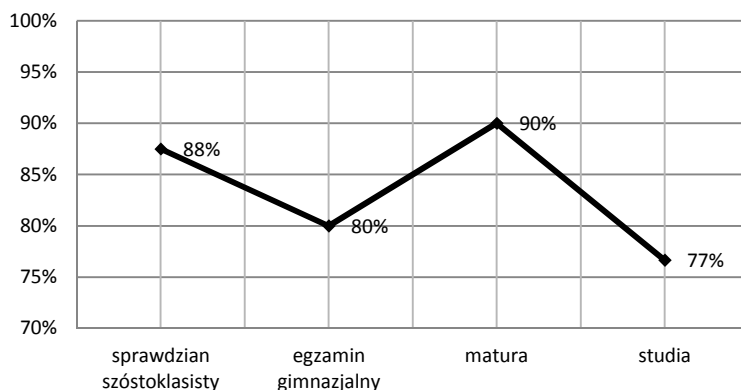
Wyniki uzyskiwane przez uczniów/studentów wykorzystamy do szacowania wartości kapitału ludzkiego wiedząc, że wiedza i umiejętności są jednym z najważniejszych jego składników.

3. Opis eksperymentu

W próbie badawczej wykorzystano wyniki 156 uczniów/studentów. Najpierw zostały przygotowane i opracowane dane dotyczące wyników uzyskanych na I roku studiów przez studentów Państwowej Wyższej Szkoły Informatyki i Przedsiębiorczości w Łomży w roku akademickim 2009/2010. Następnie dane te zostały scalone z wynikami uzyskanymi przez tych studentów we wcześniejszych poziomach edukacji. W tym momencie niezbędna była pomoc OKE w Łomży, która przygotowała dane, dla poszczególnych studentów, podając ich wyniki z egzaminu maturalnego, gimnazjalnego i szóstoklasisty. Należy podkreślić fakt, że byli to studenci po pierwszym roku studiów, którzy rocznikowo, jako pierwsi przeszli przez cały system egzaminowania zewnętrznego w Polsce.

Wszystkie wyniki egzaminów analizowane są według jednego wzoru, tzn. wynik uzyskany przez danego ucznia/studenta jest przeliczany na udział procentowy w wyniku maksymalnym. Na przykład: uczeń Jan Kowalski (JK) na sprawdzianie szóstoklasisty uzyskał 35 punktów na 40 punktów maksymalnie możliwych, czyli ten pierwszy egzamin zdał na 88%. Na egzaminie gimnazjalnym ten sam uczeń otrzymał w sumie 80 punktów (45 punktów z części

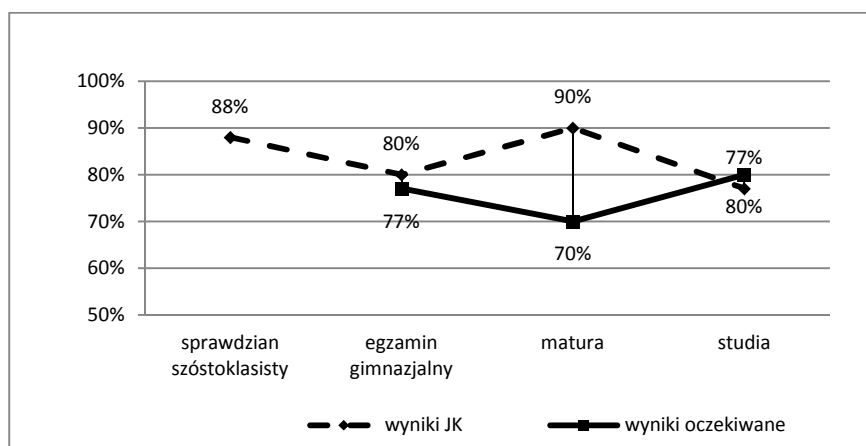
humanistycznej i 35 punktów z części matematyczno-przyrodniczej) na 100 punktów możliwych, czyli zdał na 80%. Na maturze natomiast JK otrzymał w sumie 153 punkty (z języka polskiego 64 punkty, z języka obcego 45 punktów, z przedmiotu obowiązkowego 44 punkty) na 170 maksymalnie możliwych do uzyskania punktów, czyli zdał na 90%. Natomiast na I roku studiów ten student z możliwych do uzyskania 300 punktów zdobył w sumie 230, czyli zdał na 77% (rys. 3).



Rys. 3. Wyniki JK na kolejnych egzaminach w %

W pomiarze wartości kapitału ludzkiego będziemy nie tylko wykorzystywali zebrane wyniki od szkoły podstawowej po studia, ale będziemy je również porównywać z wynikami uzyskiwanym przez uczniów/studentów całej badanej grupy. Sam wynik surowy, punktowy lub procentowy danego ucznia/studenta, nie pokazuje skali pracy, wysiłku jaki wniósł on w budowanie swojego kapitału ludzkiego, dopiero porównanie z wynikami przewidywanymi dla grupy, koleżanek i kolegów danej osoby, pokazuje nam właściwy przyrost jego kapitału ludzkiego.

Rozważmy jeszcze raz przykład JK opisany powyżej, z tym że teraz porównamy jego osiągnięcia z wynikami grupy 156 jego rówieśników (rys. 4). Wyniki analizowane na rys. 4 to rezultat regresji liniowej, o którym mówimy w punkcie 3.1.



Rys. 4. Przyrosty wartości kapitału ludzkiego JK

Z analizy tego prostego przykładu wynika, że JK egzamin gimnazjalny napisał na 80%, natomiast wynik oczekiwany wyniósł 77%. Wynik ucznia na egzaminie okazał się lepszy niż wynik oczekiwany, który został wyliczony na podstawie wyników egzaminacyjnych badanej grupy. Wynik wyższy niż wynik oczekiwany wskazuje na przyrost wartości kapitału ludzkiego tego ucznia, w danym okresie nauki szkolnej. Janek Kowalski lepiej, bardziej efektywniej wykorzystał czas nauki w szkole, lepiej niż jego koledzy pracował on nad budowaniem swojego kapitału ludzkiego. W szkole średniej widzimy znowu wyraźny wzrost powyżej oczekiwań (wynik maturalny to 90%, wynik oczekiwany to 70%). Bardzo dobry wynik maturalny wskazuje na dalsze inwestowanie we własny kapitał ludzki JK. Na I roku studiów JK zdał egzaminy na 77% (wartość oczekiwana 80%), czyli jego poziom przyrost kapitału ludzkiego był niższy w porównaniu do przyrostów kapitału ludzkiego pozostałych w grupie studentów I roku.

3.1. Szacowanie wartości oczekiwanych kapitału ludzkiego za pomocą regresji liniowej

Celem badania jest stwierdzenie, czy między badanymi zmiennymi (wynikami na wejściu i wynikami na wyjściu) zachodzi jakaś zależność, a jeśli tak, to jaka jest jej siła, kształt i kierunek. Inaczej rzecz ujmując pytamy jaka jest między nimi współzależność. Będziemy zatem ustalać jak zmieni się, średnio biorąc, wynik egzaminacyjny na wyjściu Y w zależności od wartości wyniku

egzaminacyjnego na wejściu X . Badanie korelacyjne ma sens jedynie wtedy, gdy założymy, że między zmiennymi istnieje więź przyczynowo-skutkowa, czyli że wyniki wyższego szczebla nauki są skutkiem wypracowanego wcześniej wyniku niższego szczebla – przyczyny.

Celem badań jest zatem analiza jakościowa i ilościowa związku (współzależności) między przyczyną i skutkiem, między wynikami uzyskiwanymi przez uczniów/studentów na kolejnych szczeblach edukacji. Zależność ta będzie oczywiście jednostronna, gdyż trudno jest sobie wyobrazić, że wynik uzyskany na wyższym szczeblu mógłby wpłynąć na wynik niższego szczebla edukacji. Zawsze będzie to zależność jednokierunkowa pomiędzy wynikiem niższego szczebla (wcześniejszym), a kolejnym, czyli wyższego szczebla. Kierunek zależności korelacyjnej jest od wyniku szóstoklasisty do egzaminu gimnazjalisty i kolejno od wyniku gimnazjalnego do maturalnego, i dalej od maturalnego do wyniku uzyskanego na pierwszym roku studiów

$$v(KL, S_i, t_{sp}) \rightarrow v(KL, S_i, t_{gim}) \rightarrow v(KL, S_i, t_{mat}) \rightarrow v(KL, S_i, t_{st})$$

gdzie:

- $v(KL, S_i, t_{sp})$ – wartość KL ucznia na koniec szkoły podstawowej i równocześnie początkowa wartość KL ucznia w gimnazjum,
- $v(KL, S_i, t_{gim})$ – wartość KL ucznia na koniec gimnazjum i początkowa wartość KL ucznia w szkole średniej,
- $v(KL, S_i, t_{mat})$ – wartość KL ucznia na koniec szkoły średniej i początkowa wartość KL studenta na studiach,
- $v(KL, S_i, t_{st})$ – wartość KL studenta po pierwszym roku studiów.

W celu zbadania stopnia zależności pomiędzy zmiennymi zależnymi i niezależnymi posłużymy się współczynnikiem korelacji i funkcją regresji, czyli określimy siłę tej zależności oraz jakiej zmiany zmiennej zależnej należy oczekiwać przy zmianie o jeden zmiennej niezależnej.

Współczynnik korelacji liniowej Pearsona r_{xy} jest miernikiem siły związku liniowego między dwiema cechami mierzalnymi, czyli jednostkowym przyrostom jednej zmiennej (przyczyny – wyniki uzyskane na niższym szczeblu) towarzyszy, średnio, stały przyrost drugiej zmiennej (skutki – wyniki uzyskane na wyższym szczeblu). Współczynnik korelacji liniowej Pearsona jest ściśle powiązany z kowariancją i średnimi odchyleniami od średniej arytmetycznej. Jak wiadomo ze statystyki, kowariancja jest średnią arytmetyczną iloczynu odchyleń wartości X i Y od ich średnich arytmetycznych i wyznacza się według wzoru

$$cov(x, y) = cov(y, x) = \frac{1}{n} \sum (x - \bar{x})(y_i - \bar{y})$$

Kowariancja pomiędzy wynikami na poszczególnych stopniach edukacji dla badanej populacji wynosi

$$\text{cov}(x_{sp}, y_{gim}) = 0,011669$$

$$\text{cov}(x_{gim}, y_{mat}) = 0,009119$$

$$\text{cov}(x_{mat}, y_{st}) = 0,006341414$$

gdzie:

x_{sp} – wyniki uzyskane przez uczniów na egzaminie szóstoklasisty,

$x_{gim} = y_{gim}$ – wyniki uzyskane przez uczniów na egzaminie gim.,

$x_{mat} = y_{mat}$ – wyniki uzyskane przez uczniów na maturze,

y_{st} – wyniki uzyskane przez studentów na I roku studiów.

Kowariancja obrazuje następujące informacje o związku korelacyjnym: $\text{cov}(x, y) = 0$ – brak zależności korelacyjnej, $\text{cov}(x, y) < 0$ – ujemna zależność korelacyjna, $\text{cov}(x, y) > 0$ – dodatnia zależność korelacyjna. W naszym przypadku dane wykazują dodatni związek funkcyjny pomiędzy wynikami na wejściu a wynikami na wyjściu. Zależność wyników uzyskiwanych na wyższym szczeblu edukacji pozostają w dodatniej zależności korelacyjnej z wynikami na niższym szczeblu edukacji.

Współczynnik korelacji liniowej Pearsona wyznacza natężenie i kierunek współzależności liniowej dwóch zmiennych losowych X i Y .

$$r_{xy} = r_{yx} = \frac{\text{cov}(x, y)}{s(x)s(y)}$$

gdzie $\text{cov}(x, y)$ jest kowariancją pomiędzy zmiennymi X i Y , $s(x)$ i $s(y)$ to odpowiednio odchylenia standardowe tych zmiennych.

Dla badanej populacji współczynniki korelacji liniowej Pearsona odpowiednio wynoszą

$$r_{x_{sp}y_{gim}} = 0,664783714$$

$$r_{x_{gim}y_{mat}} = 0,537098584$$

$$r_{x_{mat}y_{st}} = 0,498154746$$

Współczynnik korelacji liniowej Pearsona jest miarą unormowaną i przyjmuje wartości z przedziału od -1 do $+1$. Orientacyjnie przyjmuje się, że korelacja między dwiema cechami jest niewyraźna, jeżeli $r_{xy} \leq 0,3$, średnia, jeżeli $0,3 < r_{xy} \leq 0,5$ i wyraźna, jeżeli $r_{xy} > 0,5$. Z naszej analizy wynika, że w dwóch przypadkach ta zależność jest dodatnia i wyraźna, największa dla zależności wyników gimnazjalnych od wyników egzaminu szóstoklasisty, a średnią zależność odnotowano przy relacji wyników uzyskanych przez studentów na I roku studiów a wynikami maturalnymi.

Kwadrat współczynnika korelacji nazywamy współczynnikiem determinacji (określoności). Współczynnik determinacji informuje o tym, jaka część zmian zmiennej objaśnianej (skutek – wyjście WTP) jest wyjaśniona przez zmiany zmiennej objaśniającej (przyczyna – wejście WTP)

$$\begin{aligned} (r_{x_{sp}y_{gim}})^2 &= 0,441937386 \\ (r_{x_{gim}y_{mat}})^2 &= 0,288474889 \\ (r_{x_{mat}y_{st}})^2 &= 0,248158151 \end{aligned}$$

Oznacza to, że wyniki gimnazjalne w 44% są wyjaśnione (objaśnione) na podstawie wyników szóstoklasisty, w 28% maturalne objaśnione gimnazjalnymi i odpowiedni w tylko 24% wyników z I roku studiów jest objaśniona wynikami maturalnymi.

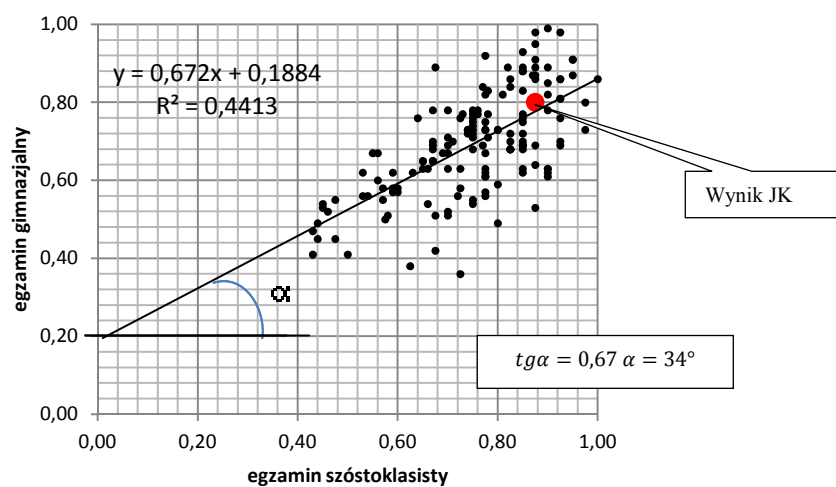
Funkcja regresji to analityczny wyraz przyporządkowania średnich wartości zmiennej objaśnianej (zależnej) konkretnym wartościom zmiennych objaśniających (niezależnych). My funkcję regresji wykorzystamy do graficznej analizy zależności wyników egzaminów wyższego szczebla od uzyskanych wyników na niższym szczeblu. W każdej funkcji regresji można wyróżnić następujące elementy składowe: zmienną objaśnianą (zależną), zmienną objaśniającą, typ związku funkcyjnego, parametry strukturalne oraz składnik losowy. Ogólna postać funkcji regresji liniowej: $y = ax + b$, gdzie: a, b – parametry liniowej funkcji regresji Y względem X .

Szacowanie parametrów strukturalnych funkcji regresji odbywa się metodą najmniejszych kwadratów. W przypadku badanej populacji otrzymujemy

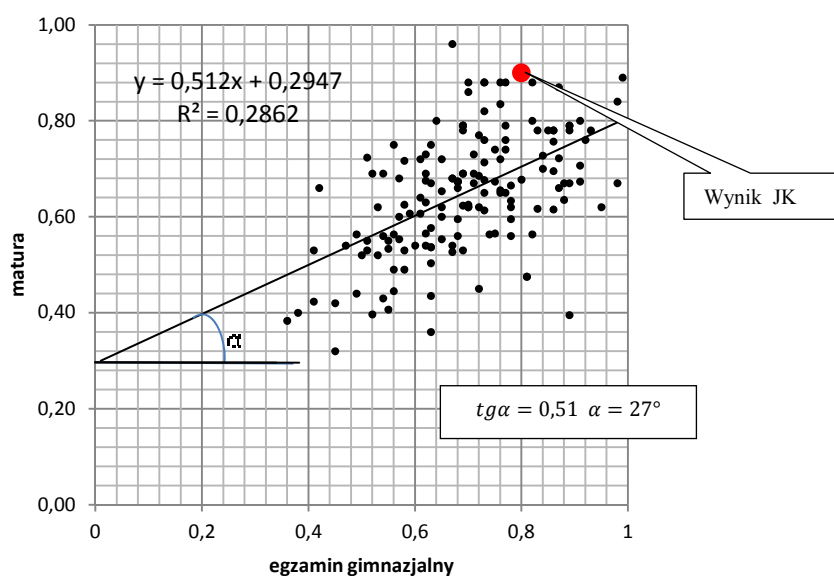
$$\begin{aligned} y &= 0,6734x + 0,1876 & a &= 0,6734, b = 0,1876 \\ y &= 0,5128x + 0,2939 & a &= 0,5128, b = 0,2939 \\ y &= 0,3912x + 0,4067 & a &= 0,3912, b = 0,4067 \end{aligned}$$

Parametry a i b nazywamy współczynnikami regresji liniowej. Parametr a odpowiada na pytanie, jaki jest przeciętny przyrost wartości zmiennej zależnej na jednostkę przyrostu wartości zmiennej niezależnej, natomiast b jest wartością stałą.

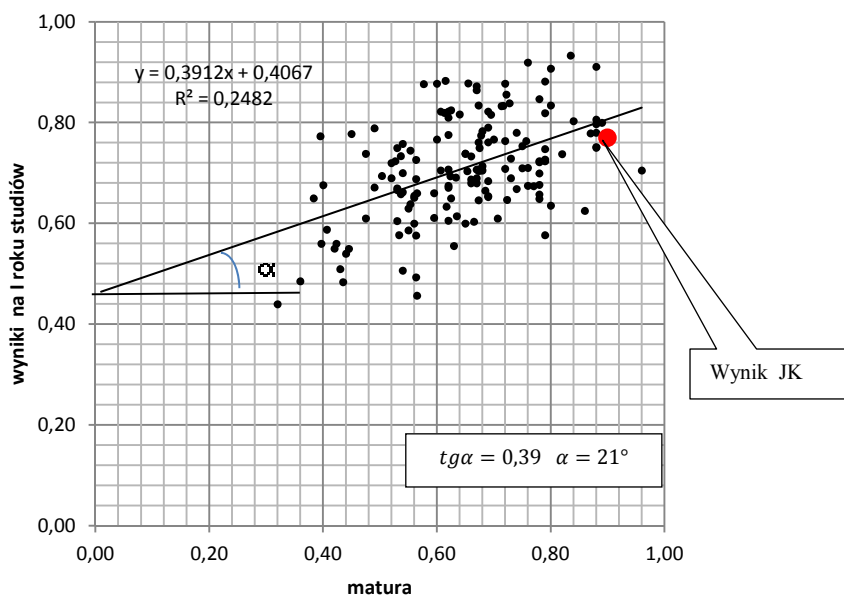
W pierwszej zależności, gdzie zmienną zależną jest wynik gimnazjalny, a zmienną niezależną wynik szóstoklasisty, $a = 0,6734$, co oznacza, że wzrost oceny szóstoklasisty o 1% powoduje średni wzrost oceny gimnazjalnej o około 0,67% (rys. 5). W drugiej zależności, gdzie zmienną zależną jest wynik maturalny, a zmienną niezależną wynik gimnazjalny, $a = 0,5128$, co oznacza, że wzrost oceny gimnazjalnej o 1% powoduje średni wzrost oceny maturalnej o około 0,51% (rys. 6). W trzeciej zależności, gdzie zmienną zależną jest wynik I roku studiów, a zmienną niezależną wynik maturalny, $a = 0,3912$, co oznacza, że wzrost oceny maturalnej o 1% powoduje średni wzrost oceny na I roku studiów o około 0,39% (rys. 7).



Rys. 5. Wykres zależności wyniku egzaminu gimnazjalnego od wyniku szóstoklasisty



Rys. 6. Wykres zależności wyniku egzaminu maturalnego od wyniku gimnazjalnego

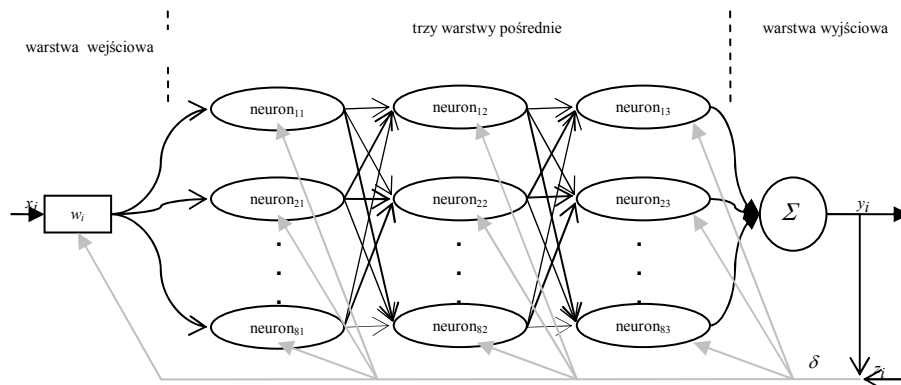


Rys. 7. Wykres zależności wyniku na I roku studiów od wyniku maturalnego

Teraz przyjrzyjmy się dokładnie kątom nachylenia liniowych regresji do osi X (rys. 5-7). Kąt α jest kątem nachylenia prostej regresji Y względem osi odciętych X , a $\operatorname{tg} \alpha = a$. W naszym przypadku kąt nachylenia prostej regresji względem osi odciętych X maleje, wraz z przejściem z jednego szczebla edukacji na kolejny. Kąt nachylenia regresji wyniku gimnazjalnego od wyniku szóstoklasisty wynosi około 34° i kolejno kąt nachylenia regresji wyniku maturalnego od gimnazjalnego wynosi 27° , natomiast kąt nachylenia regresji wyników na I roku studiów względem wyniku maturalnego wynosi 21° . Oznacza to, że wraz z przejściem na wyższy poziom edukacji (na kolejny odcinek WTP) obniżają się wyniki egzaminacyjne (wyjściowe) bardzo dobrych uczniów/studentów.

3.2. Szacowanie wartości oczekiwanej kapitału ludzkiego za pomocą sztucznej sieci neuronowej

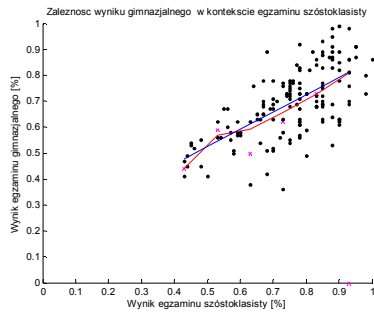
W kolejnym kroku naszych badań zaproponowano model sztucznej sieci neuronowej do szacowania wartości kapitału ludzkiego. W obliczeniach zastosowano wielowarstwową jednokierunkową sieć neuronową, której model przedstawiono na rys. 8.

Rys. 8. Model sieci neuronowej do szacowania $V(KL)$

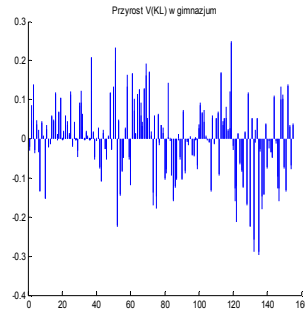
Sieć składa się z warstwy wejściowej, w której podawane są sygnały odpowiadające danym wejściowym x_i , poziom kapitału ludzkiego ucznia/studenta na wejściu danej WTP, trzech warstw pośrednich, w których sieć trenowana jest na ośmiu neuronach w każdej z nich oraz z warstwy wyjściowej z wynikiem y_i , poziom kapitału ludzkiego ucznia/studenta na wyjściu WTP. W modelu tym zastosowano w pierwszej warstwie nieliniową funkcję aktywacji (sigmoidalną), w pozostałych warstwach funkcje liniowe. Zadaniem sztucznej sieci neuronowej jest wyliczenie przyrostu wartości kapitału ludzkiego $V(KL)$ oraz narysowanie krzywej regresji – zależności wyników na kolejnym szczeblu edukacji od wyniku wcześniejszego/poprzedniego. Krzywa ta nie jest linią prostą, tak jak to było w regresji liniowej i wykazuje większe dopasowywanie się do średnich wyników egzaminacyjnych.

W procesie uczenia sieci wykorzystano metodę wstecznej propagacji błędów, której algorytm opiera się na sprowadzeniu wartości funkcji błędu sieci. Podstawą do jego obliczenia stanowił błąd średniokwadratowy MSE. Wygenerowano wykresy korelacji oraz wykresy „wydajności” (performance) w programie MATLAB 2009a. Na podstawie przeprowadzonych analiz dla wszystkich 156 uczniów/studentów w 3 zestawach testowych otrzymano 6 wykresów przedstawiających zależność pomiędzy wartościami na wyjściu (output), a wartościami przewidywanymi (target). Otrzymano tym samym krzywe regresji kawałkami liniowej. Wykresy te uzyskano dla zbioru danych uczących, walidacyjnych, testowych oraz wszystkich łącznie. Poniżej przedstawiono wynikowe wykresy dla wszystkich zestawów testowych (rys. 9-14).

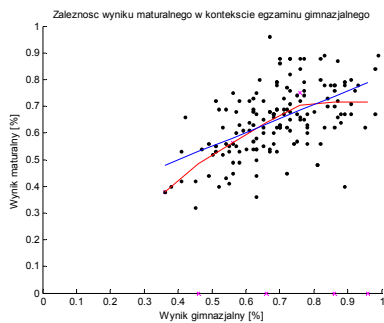
Ogólna idea polega na zastąpieniu regresji liniowej wielowarstwową siecią neuronową. Sieć na wejściu dostaje wynik niższego szczebla, a wartością oczekiwaną na wyjściu jest wynik z egzaminu wyższego szczebla edukacji. Sieć trenujemy podając jej kolejno każdą z obserwacji i pozwalając jej na takie dobranie współczynników wag, aby jak najlepiej dostosowała się do danych. Po wytrenowaniu, sieć będzie potrafiła wytypować wynik oczekiwany.



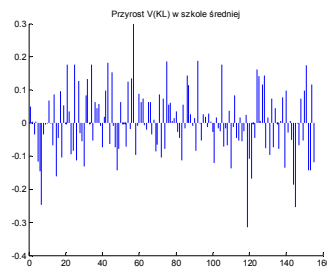
Rys. 9. Krzywa regresji dopasowania wyników gimnazjalnych do wyników szóstoklasisty wyznaczona przez SSN



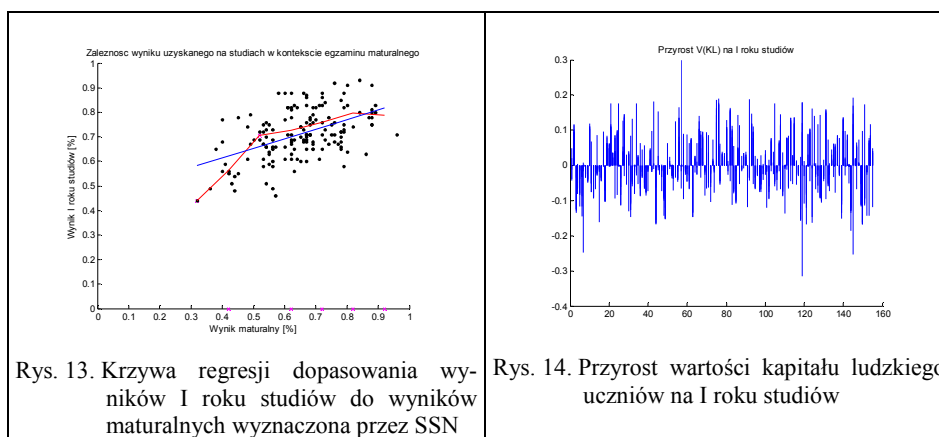
Rys. 10. Przyrost wartości kapitału ludzkiego uczniów w gimnazjum



Rys. 11. Krzywa regresji dopasowania wyników maturalnych do wyników gimnazjalnych wyznaczona przez SSN



Rys. 12. Przyrost wartości kapitału ludzkiego uczniów w szkole średniej



Rys. 13. Krzywa regresji dopasowania wyników I roku studiów do wyników maturalnych wyznaczona przez SSN

Rys. 14. Przyrost wartości kapitału ludzkiego uczniów na I roku studiów

Widać, że sieć bardziej stara się dostosować do danych, zauważalny jest ruch krzywej w kierunku średnich wartości wyników egzaminacyjnych. Nieliniowa funkcja aktywacji inaczej premiuje odległości bliższe, a inaczej dalsze. Efektem ubocznym jest to, że wynik jednak jest dużo bardziej nieprzewidywalny – w zasadzie jedynie określenie odpowiednio dużej liczby cykli uczących może nam zapewnić powtarzalność wyników, a wartość graniczna będzie różna dla różnych zbiorów danych. Ogólna tendencja została zachowana – jeśli uczeń miał dodatni przyrost wartości $V(KL)$ obliczony za pomocą regresji liniowej, to również przy użyciu sieci neuronowej wyznaczony został dodatni $V(KL)$.

Mimo to, że wytrenowana sieć neuronowa dała lepsze wyniki dopasowania wyników w poszczególnych zestawach korelacyjnych, to musimy pamiętać, że metoda wyliczania przyrostów wartości kapitału ludzkiego z wykorzystaniem SSN jest dużo bardziej skomplikowana i pracochłonna, a różnica w wynikach jest mało istotna (tabela 2).

Tabela 2

Porównanie wyników uzyskanych przez SSN i RL

Regresja wyników	MSE	RBS	$\sum V(KL)$ SSN	$\sum V(KL)$ RL
Gimnazjalne od szóstoklasisty	0,00092	0,100262	-0,00046	0,00265
Maturalne od gimnazjalnych	0,00012	0,108099	0,00964	0,00406
I rok studiów od maturalnych	0,00059	0,087256	0,05159	0,00065

Przy mniejszych błędach pracy sztucznej sieci neuronowej uzyskujemy lepsze dopasowanie krzywej regresji do wyników.

Podsumowanie

Bardzo ciekawym spostrzeżeniem jest to, że wraz z kolejnymi przejściami przez szczeble edukacji prosta regresja coraz bardziej nachyla się do osi X . Nie można tego wyraźnie zaobserwować w sieciach neuronowych, dlatego że krzywa regresji jest kawałkami liniowa. Interpretacja wyników uzyskanych przy pomocy sieci neuronowych wykorzystujących nieliniowe funkcje aktywacji jest trudnym zadaniem. Problemem jest tu brak dodatkowych danych – na wynik egzaminów składa się wiele czynników. Zmniejszenie kąta nachylenia może oznaczać, że zależność wyników na coraz wyższych szczeblach w mniejszym stopniu zależy od wyników na wejściu. Z doświadczenia autorki wynika, że nie jest to związane z faktycznym spadkiem wiedzy uczniów/studentów, wręcz przeciwnie. Na te odnotowane słabsze wyniki może wpływać wiele czynników, np. trudna interpretacja wyników egzaminów, przede wszystkim maturalnych, brak spójności w przeprowadzanych egzaminach zewnętrznych, uczenie pod egzamin, brak motywacji do lepszej pracy uczniów/studentów itp.

Mimo tych trudności, powinniśmy próbować mierzyć wartość kapitału ludzkiego oraz społecznego i tym samym wypełniać lukę między wartością rynkową a wartością księgową firmy. Nie da się zarządzać czymś, czego nie można zmierzyć. Znając wartości kapitału ludzkiego i społecznego lepiej będziemy rozumieli niefinansowe i niematerialne aktywa firmy, szkoły, uczelnie.

Literatura

1. Becker G.S. (1975). Human Capital. NBER, New York.
2. Domański S.R. (1993). Kapitał ludzki i wzrost gospodarczy. PWN, Warszawa.
3. Rutkowski L. (2009). Metody i techniki sztucznej inteligencji. PWN, Warszawa.
4. Kasperski M. (2003). Sztuczna inteligencja. Helion, Gliwice.
5. Niemierko B. (2004). Diagnostyka edukacyjna. Teoria i praktyka. Z okazji dziesiątej konferencji. Kraków.
6. Nowiński M. (1998). Prognozowanie przy użyciu sieci neuronowych. W: Metody i zastosowania badań operacyjnych, część I. Red. T. Trzaskalik. AE, Katowice.
7. Tadeusiewicz R., Gąciarz T., Borowik B., Leper B. (2007). Odkrywanie właściwości sieci neuronowych przy użyciu programów w języku C#. Wydawnictwo PAU.
8. Tadeusiewicz R. (1999). Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami. Akademicka Oficyna Wydawnicza PLJ, Warszawa.

9. Walukiewicz S. (2006). Systems Analysis of Social Capital at the Firm Level, Working Paper WP-1-2006. Systems Research Institute, Warsaw.
10. Walukiewicz S., Wiktorzak A.A. (2008). The Virtual Production Line as a New Model in Research on Education. Referat wygłoszony na TSCF Conference, Malta.
11. Walukiewicz S. (2008). Piękno liczby cztery (w naukach społecznych). IBS PAN, Warszawa.
12. Walukiewicz S. (2008). The Dimensionality of Capital and Proximity. Working Paper, System Research Institute, Warsaw.
13. Walukiewicz S. (2010). Kapitał ludzki, skrypt akademicki. IBS PAN, Warszawa.
14. Wiktorzak A.A. (2008). Budowa kapitału ludzkiego i społecznego na przykładzie szkoły ponadgimnazjalnej. Uczenie się i egzamin w oczach nauczyciela. XIV KKDE, Opole.
15. Wiktorzak A.A. (2009). Zastosowanie sztucznych sieci neuronowych do pomiaru kapitału ludzkiego. W: Analiza Systemowa w finansach i zarządzaniu. Wybrane problemy. Tom 11. Red. J. Hołubca. IBS PAN, Warszawa.
16. Wiktorzak A.A. (2009). Wirtualna Taśma Produkcyjna w badaniach edukacyjnych. XV Konferencja Diagnostyki Edukacyjnej. Kielce.
17. Wiktorzak A.A. (2009). Obliczenia inteligentne a kapitał społeczny szkoły. XV Konferencja Diagnostyki Edukacyjnej. Kielce.

METHODS FOR MEASURING ACHIEVEMENT ACROSS EDUCATION LEVELS

Summary

This paper presents a thesis that human capital is the most general measure of student achievement across learning stages. We discuss methods to measure/estimate its value, provide a definition of human capital of the student and describe stages of building its stocks. We present role of Polish educational system in it and discuss two key methods of HC value assessment: by linear regression or by artificial neural networks. We present results of extensive experiment whose principal objective was to compare practicality of either. The experiment involved 156 students whose performance was measured at each education level – from primary through secondary school through high school to the first year of college.

In final part we offer conclusions regarding utility of this type of research in assessing overall value of a company, school or university.