

PRACE NAUKOWE

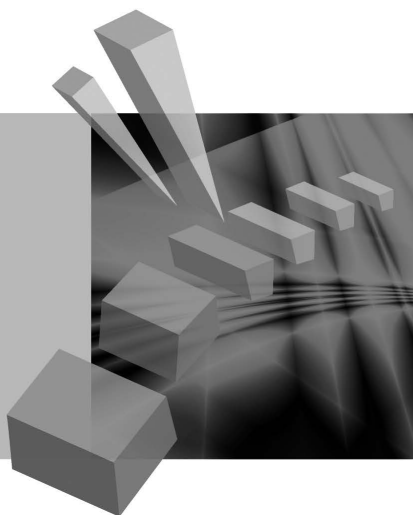
Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

238

Zastosowania badań operacyjnych Zarządzanie projektami, decyzje finansowe, logistyka



Redaktor naukowy

Ewa Konarzewska-Gubała



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2011

Recenzenci: Stefan Grzesiak, Donata Kopańska-Bródka, Wojciech Sikora,
Józef Stawicki, Tomasz Szapiro, Tadeusz Trzaskalik

Redaktor Wydawnictwa: Elżbieta Kożuchowska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

www.ibuk.pl, www.ebscohost.com,

The Central and Eastern European Online Library www.ceeol.com,

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się

na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Kopiowanie i powielanie w jakiegokolwiek formie

wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2011

ISSN 1899-3192

ISBN 978-83-7695-195-9

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp.....	9
------------	---

Część 1. Zarządzanie projektami i innowacjami

Tomasz Błaszczyk: Świadomość i potrzeby stosowania metod badań operacyjnych w pracy polskich kierowników projektów	13
Barbara Gładysz: Metoda wyznaczania ścieżki krytycznej przedsięwzięć z rozmytymi czasami realizacji zadań	25
Marek Janczura, Dorota Kuchta: Proactive and reactive scheduling in practice.....	34
Tymon Marchwicki, Dorota Kuchta: A new method of project schedule levelling	52
Aleksandra Rutkowska, Michał Urbaniak: Harmonogramowanie projektów na podstawie charakterystyk kompetencji – wrażliwość modelu na różne aspekty liczb rozmytych	66
Jerzy Michnik: Zależności między kryteriami w wielokryterialnych modelach zarządzania innowacjami	80

Część 2. Podejmowanie decyzji finansowych

Przemysław Szufel, Tomasz Szapiro: Wielokryterialna symulacyjna ocena decyzji o finansowaniu edukacji wyższej	95
Marek Kośny: Koncepcja dominacji pierwszego i drugiego rzędu w analizie wzorca zmian w rozkładzie dochodu.....	111
Agnieszka Przybylska-Mazur: Podejmowanie decyzji monetarnych w kontekście realizacji celu inflacyjnego	120
Agata Gluzicka: Analiza ryzyka rynków finansowych w okresach gwałtownych zmian ekonomicznych	131
Ewa Michalska: Zastosowanie prawie dominacji stochastycznych w konstrukcji portfela akcji	144
Grzegorz Tarczyński: Analiza wpływu ogólnej koniunktury giełdowej i wzrostu PKB na stopy zwrotu z portfela akcji przy wykorzystaniu rozmytych modeli Markowitza.....	153

Część 3. Problemy logistyki, lokalizacji i rekrutacji

Paweł Hanczar, Michał Jakubiak: Wpływ różnych koncepcji komisjonowania na czas realizacji zamówienia w węzle logistycznym	173
Mateusz Grzesiak: Zastosowanie modelu transportowego do racjonalizacji dostaw wody w regionie	186
Piotr Wojewnik, Bogumił Kamiński, Marek Antosiewicz, Mateusz Zawisza: Model odejść klientów na rynku telekomunikacyjnym z uwzględnieniem efektów sieciowych	197
Piotr Miszczyński: Problem preselekcji kandydatów w rekrutacji masowej na przykładzie wybranego przedsiębiorstwa	211

Część 4. Pomiar dokonań, konkurencja firm, negocjacje

Marta Chudykowska, Ewa Konarzewska-Gubała: Podejście ilościowe do odwzorowania celów strategicznych w systemie pomiaru dokonań organizacji na przykładzie strategii miasta Wrocławia	231
Michał Purczyński, Paulina Dolata: Zastosowanie metody DEA do pomiaru efektywności nakładów na reklamę w przemyśle piwowarskim	246
Mateusz Zawisza, Bogumił Kamiński, Dariusz Witkowski: Konkurencja firm o różnym horyzoncie planowania w modelu Bertrand z kosztem decyzji i ograniczoną świadomością cenową klientów	263
Jakub Brzostowski: Poprawa rozwiązania negocjacyjnego w systemie <i>Nego-Manage</i> poprzez zastosowanie rozwiązania przetargowego	296

Część 5. Problemy metodologiczne

Helena Gaspars-Wieloch: Metakryterium w ciągłej wersji optymalizacji wielocelowej – analiza mankamentów metody i próba jej udoskonalenia.	313
Dorota Górecka: Porównanie wybranych metod określania wag dla kryteriów oceny wariantów decyzyjnych	333
Maria M. Kaźmierska-Zatoń: Wybrane aspekty optymalizacji prognoz kombinowanych	351
Artur Prędko: Spojrzenie na metody estymacji w modelach regresyjnych przez pryzmat programowania matematycznego	365
Jan Schneider, Dorota Kuchta: A new ranking method for fuzzy numbers and its application to the fuzzy knapsack problem	379

Summaries

Part 1. Project and innovation management

Tomasz Błaszczuk: Awareness and the need for operations research methods in the work of Polish project managers	24
Barbara Gładysz: A method for finding critical path in a project with fuzzy tasks durations	33
Marek Janczura, Dorota Kuchta: Proaktywne i reaktywne harmonogramowanie w praktyce	51
Tymon Marchwicki, Dorota Kuchta: Nowa metoda niwelacji harmonogramu projektu	64
Aleksandra Rutkowska, Michał Urbaniak: Project scheduling using fuzzy characteristics of competence – sensitivity of the model to the use of different aspects of fuzzy numbers	79
Jerzy Michnik: Dependence among criteria in multiple criteria models of innovation management	92

Part 2. Financial decision-making

Przemysław Szufel, Tomasz Szapiro: Simulation approach in multicriteria decision analysis of higher education financing policy	110
Marek Kośny: First and second-order stochastic dominance in analyses of income growth pattern	119
Agnieszka Przybylska-Mazur: Monetary policy making in context of execution of the strategy of direct inflation targeting	130
Agata Gluzicka: Analysis of risk of financial markets in periods of violent economic changes	143
Ewa Michalska: Application of almost stochastic dominance in construction of portfolio of shares	152
Grzegorz Tarczyński: Analysis of the impact of economic trends and GDP growth in the return of shares using fuzzy Markowitz models	169

Part 3. Logistics, localization and recruitment problems

Paweł Hanczar, Michał Jakubiak: Influence of different order picking concepts on the time of execution order in logistics node	185
Mateusz Grzesiak: Application of transportation model for rationalization of water supply in the region	196
Piotr Wojewnik, Bogumił Kamiński, Marek Antosiewicz, Mateusz Zawisza: Model of churn in the telecommunications market with network effects	210

Piotr Miszczyński: The problem of pre-selection of candidates in mass recruitment on the example of the chosen company.....	227
--	-----

Part 4. Performance measurement, companies competition, negotiations

Marta Chudykowska, Ewa Konarzewska-Gubała: Quantitative approach to the organization strategy mapping into the performance measurement system: case of strategy for Wrocław city	245
Michał Purczyński, Paulina Dolata: Application of Data Envelopment Analysis to measure effectiveness of advertising spendings in the brewing industry	262
Mateusz Zawisza, Bogumił Kamiński, Dariusz Witkowski: Bertrand competition with switching cost.....	295
Jakub Brzostowski: Improving negotiation outcome in the NegoManage system by the use of bargaining solution.....	309

Part 5. Methodological problems

Helena Gaspars-Wieloch: The aggregate objective function in the continuous version of the multicriteria optimization – analysis of the shortcomings of the method and attempt at improving it.....	332
Dorota Górecka: Comparison of chosen methods for determining the weights of criteria for evaluating decision variants	350
Maria M. Kaźmierska-Zatoń: Some aspects of optimizing combined forecasts.....	363
Artur Prędko: Mathematical programming perspective on estimation methods for regression models	378
Jan Schneider, Dorota Kuchta: Nowa metoda rankingowa dla liczb rozmytych i jej zastosowanie dla problemu rozmytego plecaka	389

**Piotr Wojewnik, Bogumił Kamiński,
Marek Antosiewicz, Mateusz Zawisza**
Szkoła Główna Handlowa w Warszawie

MODEL ODEJŚĆ KLIENTÓW NA RYNKU TELEKOMUNIKACYJNYM Z UWZGLĘDNIENIEM EFEKTÓW SIECIOWYCH

Streszczenie: W pracy podjęto problem odejść klientów w telekomunikacji (*churnu*). Przyjmując, że ceny za połączenia wewnątrz sieci są niższe niż poza siecią, wykazano, że wystąpią efekty zewnętrzne zmiany operatora przez klienta. Jeżeli klient opuszcza operatora, to koszty jego rozmówców mogą rosnąć, a wtedy mogą oni podążać za odchodzącym klientem – tzw. *churn indukowany*. Z punktu widzenia operatora telekomunikacyjnego pożądane jest zidentyfikowanie klientów, których odejście wywoła najwyższy churn indukowany. W pracy posłużono się analizą symulacyjną sieci modelowanej metodą Watta-Strogatza na podstawie danych empirycznych – gdzie intensywność połączeń jest reprezentowana przez mieszaninę rozkładów prawdopodobieństwa: wykładniczego i Pareto. Wykazano, że dobrym predyktorem pozwalającym na identyfikację tych klientów jest ich ważony prestiż.

Słowa kluczowe: odchodzenie klientów, efekty sieciowe, symulacje wieloagentowe.

1. Wstęp

Podstawowym problemem funkcjonowania operatorów telekomunikacyjnych na rynku jest pozyskiwanie nowych oraz utrzymanie obecnych klientów. Choć pierwsze zadanie jest naturalnym obszarem działalności operatorów, to również poświęcają oni dużo uwagi drugiemu zadaniu, ponieważ pozyskiwanie nowego klienta jest od 5 do 8 razy bardziej kosztowne niż utrzymanie obecnego klienta w sieci operatora [Kazienko i in. 2009; Yan i in. 2001]. Szacuje się, że około 25–50% bazy klientów zmienia w ciągu roku swojego dotychczasowego operatora na innego. Konkretny odsetek klientów zmieniających operatora zależy od segmentu rynku [Furnas 2003]. Poziom rotacji jest szczególnie wysoki na rynku przedpłaconym, tzw. rynku *pre-paid*, gdzie klient może zerwać kontrakt bez większych konsekwencji.

Odsetek odchodzących od operatora klientów może być obniżany za pomocą kampanii marketingowych, nakierowanych na utrzymanie dotychczasowych klientów. Przeprowadzenie takiej kampanii wymaga oszacowania prawdopodobieństwa

odejścia klienta od operatora. Dysponując danymi przekrojowymi dotyczącymi charakterystyk społeczno-ekonomicznych klientów, można zbudować modele predykcyjne odejścia klienta. Należy zwrócić uwagę, że modele takie ulegają procesowi starzenia. Dlatego istnieje potrzeba ich częstego odświeżania i automatyzacji procesu ich budowania [Datta i in. 2000].

W literaturze wyróżnia się kilka przyczyn swobodnego churnu, tj. odejścia klienta: niesatysfakcjonujący poziom obsługi klienta, niekonkurencyjna oferta aparatów telefonicznych, wysokie ceny połączeń i niski poziom pokrycia geograficznego sieci [Hadden i in. 2007; Kim, Yoon 2004]. Czynniki te reprezentowane są w modelach prognozujących churn przez indywidualne charakterystyki socjoekonomiczne klientów, trendy w długości i liczbie połączeń [Richter i in. 2010] oraz brak satysfakcji z usług operatora [Hadden i in. 2006]. Uwzględnia się również całkowity czas trwania połączeń, poziom zadłużenia z poprzednich rachunków, staż klienta w sieci [Datta i in. 2000]. W ostatnich badaniach modele churnu są rozszerzane o charakterystyki klienta z perspektywy sieci społecznych, jak np. zdolność połączeniowa (*connectivity* – np. liczba sąsiadów w sieci) oraz wzajemna łączność (np. liczba odchodzących, połączonych sąsiadów) [Dasgupta i in. 2008]. Przedstawione czynniki socjoekonomiczne będą dalej traktowane jako przyczyny migracji indywidualnych klientów, czynniki sieciowe zaś jako przyczyny churnu indukowanego (por. rozdział 3.1).

W niniejszej pracy rozważamy problem odchodzenia klientów w kontekście sieciowej struktury usług telefonicznych. W szczególności, jeśli ceny wewnątrz (dalej nazywane *on-net*) i na zewnątrz sieci (dalej nazywane *off-net*) różnią się od siebie, to pojedyncza osoba opuszczająca operatora może spowodować zwiększenie kosztów swoich znajomych, ponieważ część ich rozmów będzie naliczana już nie według cen *on-net*, ale według wyższych – *off-net*. Osoby o podwyższonych kosztach odczuwają impuls do opuszczenia dotychczasowego operatora. W związku z tym identyfikacja osób o największym wpływie na odejście innych klientów jest istotnym zadaniem dla operatora. Jednakże liczba mierników wpływu klienta na zachowanie innych klientów w sieci jest długa, a ich kalkulacja wymaga dużej mocy obliczeniowej, ich analiza zaś – dużej ilości czasu. W związku z tym sformułowano następujące hipotezy badawcze:

1. Istnieje miara pozycji jednostki w sieci, która jest istotnie predykcyjna dla problemu indukowanego churnu.

2. Istotność takiej miary nie zależy od struktury sieci.

Przedstawione hipotezy są weryfikowane w tej pracy za pomocą eksperymentu symulacyjnego. Zbudowano model wieloagentowy rynku telekomunikacyjnego i symulowano jego funkcjonowanie w obszarze zmiany operatora przez klientów przy różnych strukturach rynku: (a) różnych intensywnościach połączeń i (b) różnych gęstościach połączeń. W dotychczasowych badaniach nie stwierdzono wpływu (c) struktury cenowej i (d) liczby operatorów, w związku z tym ich badania zaniechano na korzyść większej liczby agentów [Wojewnik i in. 2011a]. Dla każdej konfigu-

racji parametrów rynkowych wymuszano odejście jednego klienta i obserwowano, czy wywoła to również odejście od operatora klientów z najbliższego otoczenia. Wreszcie badano, które charakterystyki jednostek były najbardziej predykcyjne dla churnu indukowanego. Symulacyjna konstrukcja sieci połączeń istotnie odróżnia się od standardowo spotykanego w literaturze, gdzie modele churnu budowane są dla konkretnych fizycznych sieci. Proponowane podejście istotnie zwiększa możliwość uogólniania wyników.

Istotną różnicą względem dotychczasowych badań autorów [Wojewnik i in. 2011a] jest oparcie procesu generowania sieci połączeń pomiędzy klientami na empirycznych estymacjach natężenia połączeń – pięciu mieszaninach rozkładu wykładniczego i Pareto. Niniejsze opracowanie prezentuje rozszerzenie w stosunku do wstępnych wyników badań dla losowego natężenia połączeń między klientami, zaprezentowanych w niepublikowanym raporcie [Wojewnik i in. 2011b]. W niniejszej pracy analiza ma charakter wielowymiarowy, a w pracy [Wojewnik i in. 2011b] skupiono się głównie na prezentacji konstrukcji modelu wieloagentowego.

2. Struktura modelu

Model symulacyjny wykorzystany w badaniu ma podobną strukturę jak model przedstawiony w niepublikowanym raporcie [Wojewnik i in. 2011b] z dokładnością do istotnych różnic w parametryzacji. W modelu występują dwa rodzaje agentów: (1) abonenci sieci komórkowych oraz (2) operatorzy sieci komórkowych. Zakładamy, że liczba abonentów wynosi n , przy czym pojedynczy abonent odpowiada dokładnie jednemu numerowi telefonu, liczba operatorów zaś wynosi k .

Każdy abonent korzysta z usług jednego operatora oraz posiada listę rozmówców, do których dzwoni z ustaloną okresową intensywnością połączeń (np. miesiąca). Dzięki tym informacjom można scharakteryzować abonenta jako element sieci połączeń, tzn. reprezentować połączenia agentów w postaci ważonego skierowanego grafu G , w którym wierzchołki odpowiadają abonentom, wagi krawędzi zaś odpowiadają intensywności połączeń.

Każdy abonent może policzyć swój całkowity koszt połączeń, a także potencjalny koszt połączeń, gdyby zmienił dostawcę usług. Całkowity koszt połączeń jest funkcją intensywności połączeń (reprezentowaną przez macierz $D = [d_{ij}] \in \mathbb{R}_+$) oraz cen połączeń danego operatora k ($c_{k,1}$ dla ceny *on-net* oraz $c_{k,2}$ dla ceny *off-net*). Całkowity koszt połączeń jest równy sumie intensywności połączeń *on-net* oraz *off-net*, pomnożonej odpowiednio przez cenę *on-net* oraz *off-net*:

$$\text{koszt}_i(k) = \sum_{\{j:\text{operator}(j)=k\}} d_{ij}c_{k,1} + \sum_{\{j:\text{operator}(j)\neq k\}} d_{ij}c_{k,2},$$

gdzie i oznacza numer abonenta, k numer operatora, a funkcja $\text{operator}(j)$ wskazuje operatora wybranego przez j -tego abonenta. Abonenci podejmują decyzję o zmianie operatora na podstawie analizy potencjalnych kosztów dla różnych operatorów.

Zakładamy również, że każdy abonent potrafi zidentyfikować operatorów obsługujących wszystkich swoich znajomych. To stosunkowo silne założenie można jednak uzasadnić, szczególnie w przypadku Polski, z następujących powodów. Po pierwsze, rynek telefonii komórkowej jest dość skoncentrowany. W 2010 roku czterech największych operatorów komórkowi kontrolowali ponad 95% rynku w Polsce. Po drugie, prefiksy numerów poszczególnych operatorów są powszechnie znane. Możliwość zmiany operatora z zachowaniem numeru pojawiła się w Polsce dopiero niedawno i korzysta z niej rocznie około 2–3% abonentów, przez co prefiks numeru telefonu jest dobrym predykatorem operatora. Na koniec, abonenci mogą rozpoznać operatora wybranego przez znajomego na podstawie comiesięcznych billingów (na rynku abonenckim) lub na podstawie informacji o koszcie rozmowy (na rynku przedpłaconym).

Drugi rodzaj agenta występującego w modelu, *operator*, charakteryzuje się wyłącznie obowiązującą w jego sieci taryfą połączeń, tzn. ceną dla połączeń wewnątrzsieciovych, *on-net*, oraz ceną dla połączeń pozasieciovych, *off-net*. Jako że badanie dotyczy wpływu efektów sieciowych na zjawisko *churnu* oraz krótkoterminowych efektów sieciowych, wykluczaliśmy możliwość optymalizowania taryf cenowych przez operatorów, pozostawiając ceny połączeń jako zewnętrzne, ustalone parametry symulacji.

Do modelowania sieci połączeń wykorzystujemy topologię sieci SWN (*Small World Network*), która dobrze odzwierciedla grafy połączeń dla rzeczywistych populacji. Sieć generujemy za pomocą dwuetapowej procedury zaproponowanej przez [Watts, Strogatz 1998]. Początkowo wierzchołki (abonenci) rozmieszczane są równomiernie na okręgu (na tzw. kracie, ang. *lattice network*) i łączone ze swoimi r (parametr zwany jest promieniem) najbliższymi sąsiadami. Zatem po pierwszym etapie każdy wierzchołek posiada $2r$ połączeń z wierzchołkami znajdującymi się na prawo oraz lewo od niego. Liczba ta jest jednocześnie średnim stopniem wierzchołka w sieci [Vega-Redondo 2007]. Sieć typu *lattice network* jest wysoce sklasteryzowana, co jest zgodne z rzeczywistymi sieciami telekomunikacyjnymi. W drugim kroku sieć poddawana jest losowym zmianom, w wyniku których sieć zyskuje własność *małego świata* (tzw. właściwość *małego świata*). Dla każdego wierzchołka połączenia wynikające z konstrukcji *lattice network* dzielone są na dwie grupy: połączenia lewe i prawe. Następnie połączenia prawe są z prawdopodobieństwem μ (*rewiring probability*) usuwane oraz w ich miejsce powstaje nowe połączenie z innym losowo wybranym wierzchołkiem. Im wyższa wartość parametru *rewiring probability*, tym sieć jest bardziej zbliżona do sieci losowej niż do *lattice network*.

Parametry sieci μ oraz r zostały wyznaczone na podstawie danych anonimowego operatora telekomunikacyjnego. Na podstawie teoretycznego rozkładu prawdopodobieństwa danego w [Barrat, Weigt 2000] dobieramy takie wartości dla parametrów μ oraz r , by zminimalizować sumę kwadratów różnicy pomiędzy teoretycznym i empirycznym rozkładem rozmiaru sąsiedztwa wierzchołków. Minimalizowana funkcja ma następującą postać:

$$objective(\mu, r) = \sum_{i=1}^n (f_{\mu, r}(i) - f_{emp}(i))^2 \rightarrow \min,$$

gdzie $f_{\mu, r}(i)$ jest teoretyczną liczbą klientów, którzy posiadają dokładnie i sąsiadów dla sieci o parametrach μ oraz r . Wielkość tę porównujemy do empirycznej częstości $f_{emp}(i)$ wyznaczonej dla rzeczywistych danych. Wyestymowane w ten sposób parametry wynoszą odpowiednio $\mu = 1,00$ oraz $r = 3$. Ponieważ $\mu = 1,00$ oznacza w praktyce graf losowy, to w dalszej części pracy przyjmujemy prawdopodobieństwa zmiany połączenia na poziomie $\mu = 0,75$.

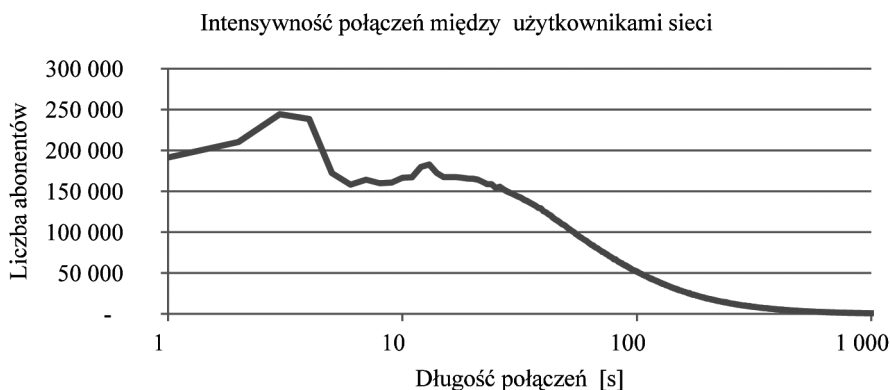
Tabela 1. Charakterystyki empirycznego rozkładu rozmiaru sąsiedztwa (liczby sąsiadów)

Min.	Kwartył 1	Mediana	Średnia	Kwartył 3	Max	Odch. std.
1	3	5	7,475	9	69	7,83

Źródło: [Wojewnik i in. 2011a].

W tabeli 1 przedstawiono charakterystykę empirycznego rozkładu rozmiaru sąsiedztwa. Średni rozmiar sąsiedztwa wynosi 7,475, co odpowiada wielkości promienia r równego ok. 3 (ze względu na medianę) lub 4 (ze względu na średnią), ponieważ średni rozmiar sąsiedztwa w *Small World Network* wynosi $2r$.

Intensywność połączeń pomiędzy dwoma abonentami mierzona jest w sekundach. Rozkład liczby połączeń został oszacowany na podstawie 4 dni połączeń (od czwartku do niedzieli) jednego z polskich operatorów telefonii mobilnej. Zanotowano ok. 20 milionów unikalnych numerów korzystających z sieci (por. rys. 1).



Rys. 1. Empiryczny rozkład długości połączeń między dwoma użytkownikami sieci

Źródło: [Wojewnik i in. 2011b].

Silna prawostronna asymetria rozkładu wskazuje na to, że dobrym przybliżeniem rozkładu intensywności połączeń będzie rozkład wykładniczy. Jednocześnie kształt rozkładu po logarytmizacji liczby połączeń wskazuje, że obserwowany ogon może być jeszcze grubszy niż w rozkładzie wykładniczym i podlega prawu skalowania (*power-law*).

Tabela 2. Charakterystyki oszacowanych rozkładów opisujących długość połączeń między dwoma dowolnymi klientami sieci telekomunikacyjnej

Mieszanina	Punkt odcięcia n^*	Exp		Pareto		
		Udział w mix	λ	Udział w mix	x_{min}	α
<i>mix1</i>	+inf	100,00%	0,0088	0,00%		
<i>mix2</i>	5890	99,96%	0,0088	0,04%	5890	3,9452
<i>mix3</i>	500	93,21%	0,0108	6,79%	500	2,6882
<i>mix4</i>	50	41,41%	0,043	58,59%	50	1,8505
<i>mix5</i>	1	0,00%		100,00%	1	1,2413

Źródło: [Wojewnik i in. 2011b].

W związku z tym do modelowania intensywności połączeń wykorzystano mieszaninę rozkładu wykładniczego, za pomocą którego modelowano główną masę prawdopodobieństwa, oraz rozkładu Pareto, za pomocą którego modelowano ogon rozkładu. Funkcja gęstości prawdopodobieństwa rozkładu wykładniczego dana jest wzorem:

$$f(x) = \lambda \cdot \exp(-\lambda \cdot x),$$

gdzie parametr $\lambda > 0$, jak również funkcja zdefiniowana jest na wartościach nieujemnych $x \in [0, +\infty)$. Funkcja gęstości prawdopodobieństwa rozkładu Pareto dana jest wzorem:

$$f(x) = k \cdot x_m^{-k} \cdot x^{-(k+1)},$$

gdzie parametry $k > 0$, $x_m > 0$, jak również funkcja zdefiniowana jest na wartościach $x \in [x_m, +\infty)$.

W tabeli 2 przedstawiono parametry oszacowanych Metodą Najwyższej Wiarygodności rozkładów w zależności od punktu odcięcia n^* .

3. Wyniki symulacji

W niniejszym rozdziale przedstawiona będzie procedura symulacyjnej analizy efektów sieciowych w churnie telekomunikacyjnym oraz wyniki uzyskane w toku tej analizy.

W niepublikowanym raporcie [Wojewnik i in. 2011b] podjęto próbę oceny, na ile wyniki badania symulacyjnego są odporne na zmiany następujących parametrów: *rewiring probability*, liczba operatorów oraz stosunek ceny połączenia *off-net* do ceny *on-net*. W tym celu przeprowadzono osobne symulacje dla różnych wartości wymienionych parametrów. Okazało się, że poszczególne parametry nie miały istotnego wpływu na uzyskiwane wartości churnu. Dlatego też w poniższym badaniu do symulacji wzięto średnie wartości powyższych parametrów oraz skoncentrowano się na zbadaniu wpływu rozkładu intensywności połączeń na wielkość churnu.

Dodatkowo w stosunku do badań [Wojewnik i in. 2011a] i [Wojewnik i in. 2011b] w niniejszej pracy znacznie zwiększamy wielkość symulacji z 3000 do 10 000 abonentów. Zmiana ta ma na celu umożliwienie empirycznego wystąpienia efektu *grubego ogona* w rozkładzie długości połączeń pomiędzy klientami.

3.1. Procedura symulacyjna

Symulacja komputerowa została przeprowadzona dla piętnastu konfiguracji parametrów sterujących procesem generowania sieci. W każdej konfiguracji 10-krotnie wygenerowano sieć połączeń i przeprowadzono symulację ich działania.

W symulacjach przyjęto następujące wartości parametrów modelu:

- liczba abonentów sieci komórkowej $n = 10\ 000$,
- liczba operatorów sieci komórkowej $k = 4$,
- stosunek ceny połączeń pozasieciowych oraz wewnątrzsieciowych $c_{i,1}/c_{i,2} = 2$,
- prawdopodobieństwo zmiany połączenia w grafie w fazie generowania połączeń sieci małego świata SWN (*rewiring probability*) $\mu = 0,75$.

Parametry różnicujące poszczególne konfiguracje sieci to promień sieci $r \in \{3,4,5\}$ oraz rozkład prawdopodobieństwa intensywności połączeń pomiędzy abonentami. Wykorzystane rozkłady są mieszaniną rozkładów wykładniczego oraz Pareto o różnych parametrach (por. tab. 2).

Pojedyncza symulacja ma następujący przebieg:

1. Losowe wygenerowanie sieci $n = 10\ 000$ klientów, w której promień sieci wynosi $r, r \in \{3,4,5\}$, prawdopodobieństwo zmiany połączeń wynosi $\mu = 0,75$, a rozkład intensywności połączeń jest zgodny z tab. 2.

2. Wyznaczenie charakterystyk sieciowych pojedynczych klientów (por. tab. 3).

3. Losowanie operatorów $k = 4$ dla poszczególnych klientów sieci i indywidualna optymalizacja tego wyboru ze względu na koszty ponoszone przez pojedynczego klienta.

4. Wymuszenie jednostkowego churnu dla poszczególnych klientów i optymalizacja wyboru operatora u wybranych innych klientów.

5. Ocena wielkości churnu indukowanego.

Procedura symulacyjna rozpoczyna się od wygenerowania sztucznej sieci typu *Small World Network* (SWN) z 10 000 klientów. Każdy z nich jest dalej charakteryzowany przez wybrane miary sieciowe.

Tabela 3. Charakterystyki sieciowe pojedynczych klientów w sieci telekomunikacyjnej

Symbol	Miara sieciowa	Interpretacja
DEG	Stopień	Liczba abonentów, z którymi abonent utrzymuje połączenia
PRE	Prestiż	Liczba abonentów wykonujących połączenia do abonenta
WPR	Ważony Prestiż	Intensywność z abonentem (liczba lub długość rozmów)
LCC	Lokalny Współczynnik Grupowania	Stopień powiązania znajomych danego abonenta (liczba połączeń między znajomymi klienta w stosunku do ich liczby)

Źródło: opracowanie własne.

Przedstawione miary DEG, PRE i WPR charakteryzują siłę powiązań pomiędzy klientami operatora telekomunikacyjnego. W szczególności miara *Stopień* (DEG) opisuje wielkość najbliższego sąsiedztwa i -tego abonenta poprzez liczbę abonentów, z którymi rozważany abonent utrzymuje kontakty telefoniczne. Liczbę abonentów, którzy dzwonią do i -tego abonenta, opisuje miara *Prestiż* (PRE), a *Ważony Prestiż* (WPR) jest liczbą abonentów ważoną przez liczbę wykonywanych przez nich połączeń. Miarą pozycji w sieci jest również *Lokalny Współczynnik Grupowania* (LCC) – będący ilorzem liczby połączeń pomiędzy znajomymi i -tego abonenta oraz liczby wszystkich możliwych połączeń, która wynosi $s(s - 1)$, gdzie s to liczba znajomych.

Po wygenerowaniu sieci połączeń abonenci są losowo przydzielani do operatorów i w oparciu o strukturę i ceny swoich połączeń optymalizują wybór operatora (por. rozdział 2). Optymalizacja taka następuje sekwencyjnie u kolejnych klientów, przy czym ich kolejność jest za każdym razem randomizowana. W przeciągu 1000 iteracji wybór operatora się stabilizuje.

Przy kolejnym etapie badania w zoptymalizowanej sieci klientów dokonywane są rozmaite zaburzenia. Dokładniej, dla każdego abonenta i losowo zmieniany jest operator (*churn wymuszony*). Zmiana taka może być interpretowana jako przejście powodowane przez czynniki pozasieciowe, np. niski poziom satysfakcji klienta ze standardów lub proponowanego aparatu telefonicznego. Po każdym zaburzeniu sieci przez okres 50 iteracji pozostali abonenci z prawdopodobieństwem $1/10$ optymalizują swój wybór operatora. Jeśli któryś z nich zmieni sieć, to ma miejsce *churn indukowany*.

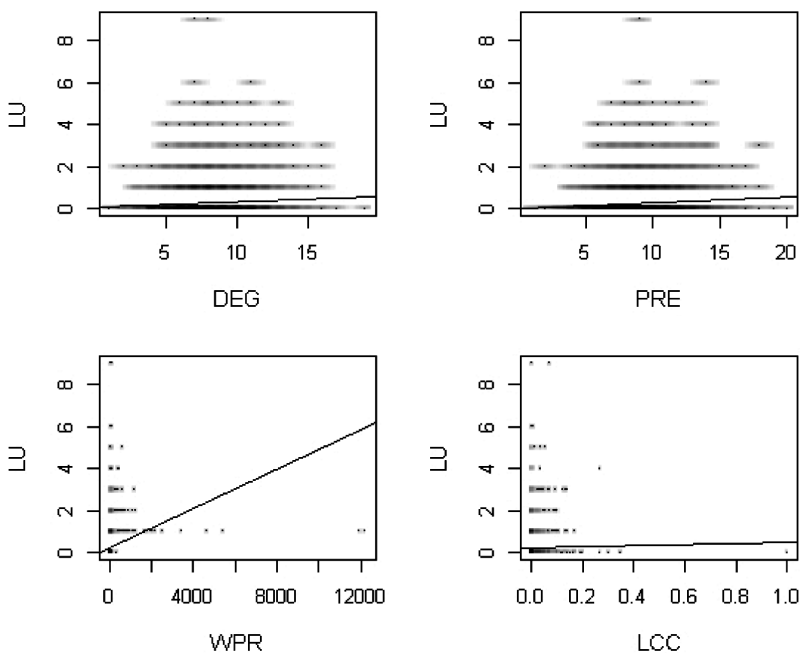
Intensywność indukowanych przejść klientów jest mierzona na dwa sposoby: LZ oraz LU. LZ oznacza liczbę abonentów, którzy dokonali jakiegokolwiek zmiany operatora, natomiast LU – jedynie liczbę ucieczek od operatora, który dotychczas obsługiwał klienta inicjującego przejścia innych.

Każdorazowo po 50 iteracjach abonenci wracają do swoich pierwotnych operatorów i dla kolejnego abonenta dokonywane jest wymuszone przejście, jego sąsiedzi zaś mają 50 iteracji na dostosowanie się.

3.2. Wyniki symulacji

Za pomocą procedury symulacyjnej opisanej w rozdziale 3.1 otrzymano zestaw charakterystyk każdego abonenta zawierający cztery miary jego pozycji w sieci oraz dwie miary jego poziomu *indukowanego churnu*. Następnie w celu zweryfikowania istotności potencjalnych predyktorów *churnu* dla każdego ze 150 zestawów danych (15 konfiguracji \times 10 powtórzeń) liczących 10 000 (liczba klientów) obserwacji zbudowano 2 liniowe modele prawdopodobieństwa, w których zmiennymi objaśnianymi są zmienne LU oraz LZ, a zmiennymi objaśniającymi są miary pozycji w sieci, tzn. DEG, PRE, WPR oraz LCC.

Uzyskane oszacowania modeli nie różnią się między sobą, ponieważ jest duża współliniowość zmiennych objaśnianych LZ oraz LU – jedynie w przypadku około 1/5 obserwacji zmienne nie są sobie równe, dla różnych obserwacji zaś różnica wynosi średnio 3,5 osoby. Wynik ten oznacza, że churn indukowany będzie występował praktycznie wyłącznie u operatora, który został opuszczony. W dalszej części zaprezentowano oszacowania parametrów modelu, w którym zmienną objaśnianą jest LU. Na rysunku 2 prezentujemy wyniki jednowymiarowej analizy zależności zmiennej LU od zmiennych objaśniających.



Rys. 2. Rozkład wartości LU względem miar pozycji poszczególnych agentów w sieci. Intensywność koloru wskazuje zagęszczenie obserwacji na danym obszarze

Źródło: opracowanie własne.

W tabeli 4 przedstawiono oszacowania parametrów modeli regresji liniowej łącznie dla wszystkich 150 symulacji (10 symulacji przy 15 konfiguracjach sieci). W ostatniej kolumnie tabeli podano, ile razy konkretna zmienna objaśniająca okazywała się istotna statystycznie dla poszczególnych parametryzacji sieci połączeń (liczba istotnych oszacowań parametrów na 150 przypadków). Dla zmiennej objaśniającej niewpływającej na zmienną objaśnianą należy spodziewać się średnio 7,5 obserwacji istotnych. Okazuje się, że miara *Ważony Prestiż* (WPR) ma największy wpływ na wielkość indukowanego churnu. Jest ona istotna statystycznie w większości symulacji. Miara *Prestiż* (PRE) również ma istotne statystycznie znaczenie przy objaśnianiu wielkości churnu. Pozostałe zmienne objaśniające wydają się nie mieć wpływu na poziom indukowanego churnu.

Tabela 4. Oszacowania parametrów liniowych modeli prawdopodobieństwa objaśniających churn

Zmienna	Współczynnik	Średni p-val	Liczba p-val < 0,05
DEG	0,02	0,45	18/150
PRE	0,18	0,02	145/150
WPR	17,88	0,00	148/150
LCC	0,22	0,49	9/150

Źródło: opracowanie własne.

W tabelach 5 i 6 przedstawiono średnie wartości współczynników regresji liniowej oraz średnie poziomy istotności współczynników regresji liniowej w zależności od parametrów symulacji – rozkładu intensywności połączenia oraz promienia sieci. Z tabel można odczytać, że podstawowe wnioski dotyczące istotności poszczególnych zmiennych są zgodne z poprzednimi i że są podobne dla różnych rozkładów prawdopodobieństwa. *Prestiż* i *Ważony Prestiż* są istotne dla znakomitej większości zastosowanej parametryzacji modelu.

Tabela 7 oraz rys. 3 przedstawiają wielkość średniego indukowanego churnu w zależności od parametrów symulacji. Poziom indukowanego churnu jest tutaj wyznaczany jako średnia liczba odejść wywołanych poprzez jeden wymuszony churn.

Okazuje się, że średni promień sieci ma istotny wpływ na intensywność badanego zjawiska. Wraz ze wzrostem liczby połączeń prawdopodobieństwo zmiany operatora pod wpływem zmiany operatora przez znajomego maleje. Z kolei sposób modelowania rozkładu prawdopodobieństwa intensywności połączeń nie ma istotnego wpływu na wielkość churnu.

Tabela 5. Średnie z 10 symulacji współczynniki regresji liniowej churnu w rozbiciu na radius i rozkład prawdopodobieństwa intensywności połączeń

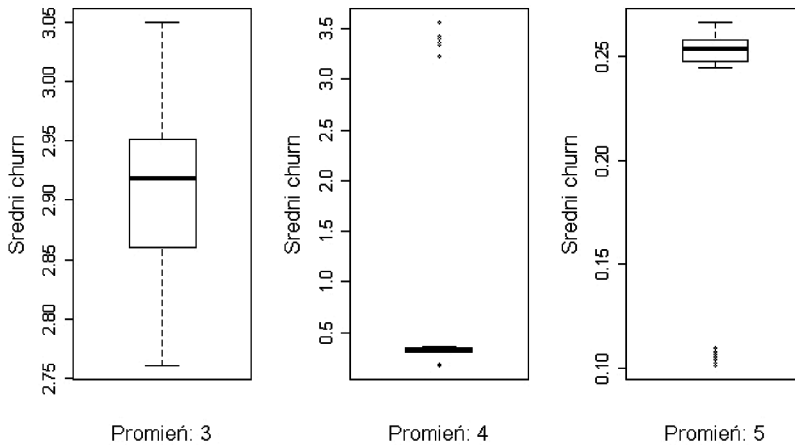
Radius	Miara sieciowa	Rozkład prawdopodobieństwa połączeń					
		1	2	3	4	5	Średnia
3	Prestiż	0,44	0,46	0,52	0,35	0,47	0,45
	Ważony Prestiż	5,46	8,56	5,86	20,66	8,07	9,72
	Wsp. Grupowania	0,72	0,50	0,54	0,79	0,33	0,58
	Stopień	0,09	0,07	0,02	0,01	0,06	0,05
4	Prestiż	0,09	0,05	0,13	0,00	0,15	0,08
	Ważony Prestiż	18,48	13,05	9,10	32,45	16,18	17,85
	Wsp. Grupowania	0,29	0,01	-0,32	-0,05	0,37	0,06
	Stopień	0,00	0,00	0,00	0,00	0,01	0,00
5	Prestiż	0,03	0,03	0,03	-0,01	0,03	0,02
	Ważony Prestiż	20,51	15,78	23,70	50,74	19,58	26,06
	Wsp. Grupowania	-0,05	0,01	0,07	-0,01	0,11	0,03
	Stopień	0,00	0,00	0,00	0,00	0,00	0,00
Prestiż, razem		0,19	0,18	0,22	0,12	0,121	0,18
Ważony Prestiż, razem		14,81	12,46	12,89	34,62	14,61	17,88
Wsp. Grupowania, razem		0,32	0,17	0,10	0,25	0,27	0,22
Stopień, razem		0,03	0,02	0,00	0,00	0,02	0,02

Źródło: opracowanie własne.

Tabela 6. Średnia z 10 symulacji istotność statystyczna współczynników regresji churnu w rozbiciu na radius i rozkład prawdopodobieństwa intensywności połączeń

Radius	Miara sieciowa	Rozkład prawdopodobieństwa połączeń					
		1	2	3	4	5	Średnia
3	Prestiż	0,00	0,00	0,00	0,00	0,00	0,00
	Ważony Prestiż	0,00	0,00	0,00	0,00	0,00	0,00
	Wsp. Grupowania	0,42	0,51	0,48	0,35	0,41	0,44
	Stopień	0,38	0,45	0,54	0,50	0,40	0,45
4	Prestiż	0,00	0,00	0,00	0,08	0,00	0,02
	Ważony Prestiż	0,00	0,00	0,02	0,00	0,00	0,00
	Wsp. Grupowania	0,54	0,40	0,51	0,37	0,52	0,47
	Stopień	0,49	0,36	0,51	0,54	0,41	0,46
5	Prestiż	0,00	0,00	0,00	0,15	0,00	0,03
	Ważony Prestiż	0,00	0,00	0,00	0,00	0,00	0,00
	Wsp. Grupowania	0,51	0,66	0,49	0,65	0,54	0,57
	Stopień	0,42	0,40	0,36	0,59	0,42	0,44
Prestiż, razem			0,00	0,00	0,07	0,00	0,02
Ważony Prestiż, razem			0,00	0,01	0,00	0,00	0,00
Wsp. Grupowania, razem			0,52	0,49	0,46	0,49	0,49
Stopień, razem			0,40	0,47	0,54	0,41	0,45

Źródło: opracowanie własne.



Rys. 3. Wykresy pudełkowe prezentują wyznaczony symulacyjnie rozkład współczynnika churnu w zależności od promienia sieci

Źródło: opracowanie własne.

Tabela 7. Średnia wielkość churnu w zależności od parametrów sieci (w procentach)

Radius	Rozkład prawdopodobieństwa połączeń				
	1	2	3	4	5
3	2,90	2,89	2,92	2,94	2,90
4	0,65	0,32	0,92	0,18	1,24
5	0,25	0,25	0,25	0,10	0,25

Źródło: opracowanie własne.

4. Uwagi końcowe

Pozyskiwanie nowych i utrzymanie dotychczasowych klientów jest warunkiem funkcjonowania przedsiębiorstw na rynku detalicznym. W badaniu podjęto problem analizy lojalności klientów w przypadku rynków sieciowych. Rozważa się wpływ pozycji sieciowej klienta na lojalność jego rozmówców, a dokładniej, czy jego przejście może wywołać churn indukowany – przejście jego sąsiadów.

Badanie miało charakter symulacyjny. Zbudowano model sieci połączeń telekomunikacyjnych, w której węzły stanowią klienci, łuki reprezentują połączenia głosowe, wagi łuków – częstość połączeń oraz w której klienci mogą być przypisani do różnych operatorów. Przyjęto, że na skutek różnic cenowych klienci mogą dokonać zmiany operatora, i obserwowano wpływ przejścia jednych klientów na decyzję o przejściu innych.

Indukowany churn badano przy wielokrotnej symulacji sieci połączeń, również przy zmianie sposobu modelowania intensywności połączeń. Spośród czterech miar pozycji sieciowej indywidualnych klientów dwie: *Prestiż* i *Ważony Prestiż* okazały się mieć istotny wpływ na wielkość churnu indukowanego przez jednostki. Ponadto globalna wielkość churnu uwarunkowana jest średnim promieniem sieci. Jeśli promień sieci wzrasta, to rośnie liczba połączeń pomiędzy abonentami, a w efekcie wpływ pojedynczych abonentów na siebie wzajemnie maleje, wielkość globalnego churnu zaś spada.

W dalszych pracach warto rozważyć inne struktury sieci, zbadać większą liczbę indywidualnych miar pozycji sieciowej oraz inne schematy przejścia i optymalizacji wyboru operatora. Należy przy tym zachować możliwie dużą liczbę agentów, aby zapewnić reprezentację również dla nietypowych agentów – których charakterystyki znajdują się w ogonach rozkładów prawdopodobieństwa poszczególnych cech, np. intensywności połączeń.

Literatura

- Barrat A., Weigt M. [2000], *On the properties of small-world network models*, "The European Physical Journal B – Condensed Matter and Complex Systems", vol. 13, no. 3, s. 547–560.
- Dasgupta K., Singh R., Viswanathan B., Chakraborty D., Mukherjea S., Nanavati A.A., Joshi A. [2008], *Social ties and their relevance to churn in mobile telecom networks*, "EDBT '08: Proceedings of the 11th international conference on Extending database technology", New York, s. 668–677.
- Datta P., Masand B., Mani D.R., Li B. [2000], *Automated cellular modeling and prediction on a large scale*, "Artificial Intelligence Review", vol. 14, s. 485–502.
- Furnas G. [2003], *Framing the wireless market. The Future of Wireless*, "WSA News:Bytes", vol. 17(11), s. 4–6.
- Hadden J., Tiwari A., Roy R., Ruta D. [2006], *Churn prediction using complaints data*, "Proc. of world academy of science, engineering, and technology", vol. 13, s. 158–163.
- Hadden J., Tiwari A., Roy R., Ruta D. [2007], *Computer assisted customer churn management: State-of-the-art and future trends*, "Computers & Operations Research", vol. 32, s. 2902–2917.
- Hastie T., Tibshirani R., Friedman J. [2008], *The Elements of Statistical Learning. Data Mining Inference and Prediction*, Springer.
- Hidalgo C., Rodriguez-Sickert C. [2007], *Persistence, Topology and Sociodemographics of a Mobile Phone Network*, "Technical report", Center for Complex Network Research, Department of Physics, University of Notre Dame.
- Katz L. [1953], *A new status index derived from sociometric analysis*, "Psychometrika", vol. 18, s. 39–43.
- Kazienko P., Brodka P., Ruta D. [2009], *The influence of customer churn and acquisition on value dynamics of social neighbourhoods*, „Lectures Notes in Computer Science”, vol. 5736/3309, s. 491–500.
- Kim H., Yoon C. [2004], *Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market*, "Telecommunications Policy" 28, s. 751–765.
- Knoke D. [2007], *Social Network Analysis in Quantitative Applications in the Social Sciences*, Sage Publications, Inc.
- Luke S., Cioffi-Revilla C., Panait L., Sullivan K., Balan G. [2005], *MASON: A multiagent simulation environment*, "SIMULATION", vol. 81 (7), s. 517–527.

- Newman M. [2003], *The structure and function of complex networks*, “SIAM Review”, vol. 45, s. 167–256.
- Onnela J., Saramaki J., Hyvonen J., Szabo G., Lazer D., Kaski K., Kertesz J., Barabasi A. [2007], *Structure and tie strengths in mobile communication networks*, “Proc. National Academy of Science USA”, vol. 104, s. 7332–7336.
- R Development Core Team: *R: A Language and Environment for Statistical Computing*, Foundation for Statistical Computing, Vienna 2010.
- Richter Y., Yom-Tov E., Slonim N. [2010], *Predicting customer churn in mobile networks through analysis of social groups*, “SIAM Int. Conf. on Data Mining”.
- Vega-Redondo F. [2007], *Complex Social Networks*, Cambridge University Press, Cambridge.
- Wasserman S., Faust K. [1994], *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.
- Watts D.J., Strogatz S.H. [1998], *Collective dynamics of ‘small-world’ networks*, “Nature”, vol. 393 (6684), s. 409–410.
- Wojewnik P., Kamiński B., Antosiewicz M., Zawisza M. [2011a], *Social-network influence on telecommunication customer attrition*, [w:] *Agent and Multi-Agent Systems: Technologies and Applications*, LNAI 6682, eds J. O’Shea, N.T. Nguyen, K. Crockett, R.J. Howlett, L.C. Jain, Springer, Heidelberg, s. 64–73.
- Wojewnik P., Kamiński B., Antosiewicz M., Zawisza M. [2011b], *Efekty sieciowe w churnie telekomunikacyjnym*, Badania Statutowe Młodych Naukowców 03/BMN/21/11, Instytut Ekonometrii, Szkoła Główna Handlowa w Warszawie.
- Yan L., Miller D.J., Mozer M.C., Wolniewicz R. [2001], *Improving prediction of customer behaviour in non-stationary environments*, “Proc. of International Joint Conference on Neural Networks, IJCNN”, vol. 3, s. 2258–2263.

MODEL OF CHURN IN THE TELECOMMUNICATIONS MARKET WITH NETWORK EFFECTS

Summary: We investigate the problem of churn in telecommunication industry considering the calls network effects. Assuming the calling party pays and the on-net/off-net prices differ we expect the network effects to appear. In particular, if an individual leaves the company, the costs of the people calling him do change. In some situations it would be fruitful for them to follow the churning person – the phenomenon called induced churn. It would be beneficial to the telco company to identify the clients potentially inducing the highest churn and our aim is to identify the factors determining the induced churn potential. For this task we investigate some network basing on empirical data. The calls network is modeled with Watts-Strogatz graph, while the calls intensity – with exponential and Pareto probability density functions.

Keywords: churn, network effects, multiagent simulation.