

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

242

Taksonomia 19.

Klasyfikacja i analiza danych – teoria i zastosowania



Redaktorzy naukowi
Krzysztof Jajuga
Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2012

Recenzenci: Eugeniusz Gatnar, Elżbieta Gołata, Tadeusz Kufel, Józef Pocięcha,
Miroslaw Szreder, Feliks Wysocki

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Tytuł sfinansowano ze środków Sekcji Klasyfikacji i Analizy Danych PTS
i Uniwersytetu Ekonomicznego we Wrocławiu

Publikacja jest dostępna na stronie www.ibuk.pl

Streszczenia opublikowanych artykułów są dostępne w międzynarodowej bazie danych
The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl>
oraz w 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](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

Kopowanie i powielanie w jakiegokolwiek formie
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2012

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM
Nakład: 320 egz.

Spis treści

Wstęp	13
Stanisława Bartosiewicz , Jeszcze raz o skutkach subiektywizmu w analizie wielowymiarowej	17
Andrzej Sokolowski , Q uniwersalna miara odległości	22
Eugeniusz Gatnar , Jakość danych w systemach statystycznych banków centralnych (na przykładzie NBP)	31
Marek Walesiak , Pomiar odległości obiektów opisanych zmiennymi mierzonymi na skali porządkowej – strategię postępowania.....	39
Krzysztof Jajuga, Marek Walesiak , XXV lat konferencji taksonomicznych – fakty i refleksje	47
Józef Pocięcha, Barbara Pawelek , Model SEM w analizie zagrożenia bankructwem przedsiębiorstw w świetle koniunktury gospodarczej – problemy teoretyczne i praktyczne	50
Paweł Lula , Uczące się systemy pozyskiwania informacji z dokumentów tekstowych	58
Ewa Roszkowska , Zastosowanie metody TOPSIS do wspomaganie procesu negocjacji.....	68
Andrzej Młodak , Sąsiedztwo obszarów przestrzennych w ujęciu fizycznym oraz społeczno-ekonomicznym – podejście taksonomiczne	76
Andrzej Bąk , Modele kategorii nieuporządkowanych w badaniach preferencji	86
Jacek Kowalewski , Zintegrowany model optymalizacji badań statystycznych.....	96
Jan Paradysz, Karolina Paradysz , Obszary bezrobocia w Polsce – problem benchmarkowy.....	106
Tomasz Szubert , W co grać, aby jak najmniej przegrać? Próba klasyfikacji systemów gry w zakładach bukmacherskich.....	116
Izabela Szamrej-Baran , Klasyfikacja krajów UE ze względu na ubóstwo energetyczne	126
Sylwia Filas-Przybył, Tomasz Klimanek, Jacek Kowalewski , Analiza dojazdów do pracy za pomocą modelu grawitacji.....	135
Marta Dziechciarz-Duda, Anna Król, Klaudia Przybysz , Minimum egzystencji a czynniki warunkujące skłonność do korzystania z pomocy społecznej. Klasyfikacja gospodarstw domowych	144
Hanna Dudek , Subiektywne skale ekwiwalentności – analiza na podstawie danych o satysfakcji z osiągniętych dochodów	153

Joanicjusz Nazarko, Ewa Chodakowska, Marta Jaročka , Segmentacja szkół wyższych metodą analizy skupień <i>versus</i> konkurencja technologiczna ustalona metodą DEA – studium komparatywne.....	163
Ewa Chodakowska , Wybrane metody klasyfikacji w konstrukcji ratingu szkół.....	173
Bartosz Soliński , Sektor energetyki odnawialnej w krajach Unii Europejskiej – klasyfikacja w świetle strategii zarządzania zmianą.....	182
Krzysztof Szwarz , Klasyfikacja powiatów województwa wielkopolskiego ze względu na sytuację demograficzną.....	192
Elżbieta Gołata, Grażyna Dehnel , Rejestry administracyjne w analizie przedsiębiorczości.....	202
Katarzyna Chudy, Marek Sobolewski, Kinga Stępień , Wykorzystanie metod taksonomicznych w prognozowaniu wskaźników rentowności banków giełdowych w Polsce.....	212
Katarzyna Dębowska , Modelowanie upadłości przedsiębiorstw przy wykorzystaniu metod dyskryminacji i regresji.....	222
Alina Bojan , Wykorzystanie metod wielowymiarowej analizy danych do identyfikacji zmiennych wpływających na atrakcyjność wybranych inwestycji.....	231
Justyna Brzezińska , Analiza logarytmiczno-liniowa w badaniu przyczyn umieralności w krajach UE.....	240
Aneta Rybicka, Bartłomiej Jefmański, Marcin Pelka , Analiza klas ukrytych w badaniach satysfakcji studentów.....	247
Bartłomiej Jefmański , Pomiar opinii respondentów z wykorzystaniem elementów teorii zbiorów rozmytych i środowiska R.....	256
Julita Stańczuk , Porównanie rezultatów wielostanowej klasyfikacji obiektów ekonomicznych z wykorzystaniem analizy dyskryminacyjnej oraz sieci neuronowych.....	265
Jerzy Krawczuk , Skuteczność metod klasyfikacji w prognozowaniu kierunku zmian indeksu giełdowego S&P500.....	275
Anna Czapkiewicz, Beata Basiura , Symulacyjne badanie wpływu zaburzeń na grupowanie szeregów czasowych na podstawie modelu Copula-GARCH.....	283
Radosław Pietrzyk , Ocena efektywności inwestycji funduszy inwestycyjnych z tytułu doboru papierów wartościowych i umiejętności wykorzystania trendów rynkowych.....	291
Aleksandra Witkowska, Marek Witkowski , Zastosowanie metody Panzara-Rosse’a do pomiaru poziomu konkurencji w sektorze banków spółdzielczych.....	306
Marcin Pelka , Podejście wielomodelowe z wykorzystaniem metody <i>boosting</i> w analizie danych symbolicznych.....	315
Justyna Wilk , Analiza porównawcza oprogramowania komputerowego w klasyfikacji danych symbolicznych.....	323

Tomasz Bartłomowicz, Justyna Wilk , Zastosowanie metod analizy danych symbolicznych w przeszukiwaniu dziedzinowych baz danych.....	333
Kamila Migdał-Najman , Propozycja hybrydowej metody grupowania opartej na sieciach samouczących	342
Dorota Rozmus , Porównanie dokładności taksonomii spektralnej oraz zagregowanych algorytmów taksonomicznych opartych na idei metody <i>bagging</i>	352
Krzysztof Najman , Grupowanie dynamiczne z wykorzystaniem samouczących się sieci GNG	361
Małgorzata Misztal , Wpływ wybranych metod uzupełniania brakujących danych na wyniki klasyfikacji obiektów z wykorzystaniem drzew klasyfikacyjnych w przypadku zbiorów danych o niewielkiej liczebności – ocena symulacyjna	370
Mariusz Kubus , Zastosowanie wstępnego uwarunkowania zmiennej objaśnianej do selekcji zmiennych.....	380
Barbara Batóg, Jacek Batóg , Wykorzystanie analizy dyskryminacyjnej do identyfikacji czynników determinujących stopę zwrotu z inwestycji na rynku kapitałowym	387
Katarzyna Wójcik, Janusz Tuchowski , Analiza porównawcza miar podobieństwa tekstów opartych na macierzy częstości i tekstów opartych na wiedzy dziedzinowej	396
Iwona Staniec , Analiza czynnikowa w identyfikacji obszarów determinujących doskonalenie systemów zarządzania w polskich organizacjach	406
Marek Lubicz, Maciej Zięba, Adam Rzechonek, Konrad Pawelczyk, Jerzy Kołodziej, Jerzy Błaszczyk , Analiza porównawcza wybranych technik eksploracji danych do klasyfikacji danych medycznych z brakującymi obserwacjami	416
Iwona Foryś , Wykorzystanie analizy log-liniowej do wyboru czynników determinujących atrakcyjność cenową mieszkań w obrocie wtórnym na przykładzie lokalnego rynku mieszkaniowego.....	426
Ewa Genge , Analiza skupień oparta na mieszankach uciętych rozkładów normalnych.....	436
Jerzy Korzeniewski , Ocena efektywności metody uśredniania zmiennych i metody Ichino selekcji zmiennych w analizie skupień	444
Andrzej Dudek , SMS – propozycja nowego algorytmu analizy skupień	451
Artur Mikulec , Metody oceny wyniku grupowania w analizie skupień.....	460
Małgorzata Machowska-Szewczyk , Algorytm klasyfikacji rozmytej dla obiektów opisanych za pomocą zmiennych symbolicznych oraz rozmytych	469
Artur Zaborski , Analiza PROFIT i jej wykorzystanie w badaniu preferencji	479
Karolina Bartos , Analiza skupień wybranych państw ze względu na strukturę wydatków konsumpcyjnych obywateli – zastosowanie sieci Kohonena	488

Barbara Batóg, Magdalena Mojsiewicz, Katarzyna Wawrzyniak , Klasyfikacja gospodarstw domowych ze względu na bodźce do zawierania umowy o ubezpieczenie z wykorzystaniem modeli zmiennych jakościowych .	496
Izabela Kurzawa , Zastosowanie modelu LA/AIDS do badania elastyczności cenowych popytu konsumpcyjnego w gospodarstwach domowych w relacji miasto–wieś	505
Aleksandra Łuczak, Feliks Wysocki , Metody porządkowania liniowego obiektów opisanych za pomocą cech metrycznych i porządkowych	513
Agnieszka Sompolska-Rzechuła , Porównanie klasycznej i pozycyjnej taksonomicznej analizy zróżnicowania jakości życia w województwie zachodniopomorskim	523
Joanna Banaś, Małgorzata Machowska-Szewczyk , Ocena intensywności wykorzystania skrzynek poczty elektronicznej za pomocą uporządkowanego modelu probitowego	532
Iwona Bąk , Segmentacja gospodarstw domowych emerytów i rencistów pod względem wydatków na rekreację i kulturę	541
Aneta Becker , Zastosowanie metody ANP do porządkowania województw Polski pod względem dynamiki wykorzystania ICT w latach 2008-2010	552
Katarzyna Dębowska , Klasyfikacja sektorów ze względu na ich kondycję finansową przy użyciu metod wielowymiarowej analizy statystycznej	562
Anna Domagała , Propozycja metody doboru zmiennych do modeli DEA (procedura kombinowanego doboru w przód).....	571
Henryk Gierszal, Karina Pawlina, Maria Urbańska , Analiza statystyczna w badaniach zapotrzebowania na usługi teleinformatyczne sieci łączności ruchomej	580
Hanna Gruchociak , Konstrukcja estymatora regresyjnego dla danych o strukturze dwupoziomowej.....	590
Tomasz Klimanek, Marcin Szymkowiak , Zastosowanie estymacji pośredniej uwzględniającej korelację przestrzenną w opisie niektórych charakterystyk rynku pracy	601
Jarosław Lira , Prognozowanie opłacalności produkcji żywca wieprzowego w Polsce	610
Christian Lis , Wykorzystanie metody klasyfikacji w ocenie konkurencyjności portów południowego Bałtyku	619
Beata Bieszk-Stolorz, Iwona Markowicz , Wykorzystanie wielomianowego modelu logitowego do oceny szansy podjęcia pracy przez bezrobotnych .	628
Lucyna Przezbórska-Skobiej, Jarosław Lira , Przestrzeń agroturystyczna Polski i ocena jej atrakcyjności.....	637
Paweł Ulman , Model rozkładu wydatków a funkcje popytu.....	646
Maria Urbańska, Tadeusz Mizera, Henryk Gierszal , Zastosowanie metod analizy statystycznej w badaniach mięczaków	655

Summaries

Stanisława Bartosiewicz , The effects of subjectivism in multivariate analysis revisited.....	21
Andrzej Sokółowski , Q universal distance measure	30
Eugeniusz Gatnar , Data quality in central banks' statistical systems (NBP example)	38
Marek Walesiak , Distance measures for ordinal data – strategies of proceedings.....	46
Krzysztof Jajuga, Marek Walesiak , XXV years of taxonomic conferences – some facts and remarks.....	49
Józef Pocięcha, Barbara Pawelek , General SEM model in researching corporate bankruptcy and business cycles – theoretical and practical problems.....	57
Paweł Lula , Learning-based systems of information extraction from textual resources	67
Ewa Roszkowska , The application of the TOPSIS method to support the negotiation process	75
Andrzej Młodak , Neighborhood of spatial areas in the physical and socio-economic context – a taxonomic approach.....	85
Andrzej Bąk , Models for unordered categories in preference analysis.....	95
Kowalewski Jacek , An integrated model of optimizing statistical surveys	105
Jan Paradysz, Karolina Paradysz , Areas of unemployment in Poland – benchmark problem	115
Tomasz Szubert , How to play to lose the least? Classification of systems in sports bets	125
Izabela Szamrej-Baran , Classification of EU member states in view of fuel poverty	134
Sylvia Filas-Przybył, Tomasz Klimanek, Jacek Kowalewski , An attempt to use the gravity model in the analysis of commuters.....	143
Marta Dziechciarz-Duda, Anna Król, Klaudia Przybysz , Subsistence minimum versus factors influencing tendency to benefit from social care. Classification of households	152
Hanna Dudek , Subjective equivalence scales – analysis based on data about satisfaction with incomes.....	162
Joanicjusz Nazarko, Ewa Chodakowska, Marta Jarocka , Segmentation of universities using cluster analysis versus technological competitors determined by the DEA method – a comparative study	172
Ewa Chodakowska , Selected methods of classification in schools' rating.....	181
Bartosz Soliński , Renewable energy sector in the European Union – classification in the light of change management strategy	191
Krzysztof Szwarc , Classification of Wielkopolska voivodeship due to the demographic situation	201

Elżbieta Gołata, Grażyna Dehnel , Administrative registers in business analysis.....	211
Katarzyna Chudy, Marek Sobolewski, Kinga Stępień , Application of taxonomic methods in forecasting the profitability ratios of listed banks in Poland.....	221
Katarzyna Dębowska , Modeling bankruptcy of firms by using discrimination and regression methods.....	230
Alina Bojan , Identification of variables which influence attractiveness of given investments with the usage of multivariate analysis.....	239
Justyna Brzezińska , Log-linear analysis in the study of mortality in EU.....	246
Aneta Rybicka, Bartłomiej Jefmański, Marcin Pelka , Latent class analysis in student satisfaction surveys.....	254
Bartłomiej Jefmański , The respondent's opinions measurement in the R program with an application of fuzzy sets theory.....	264
Julita Stańczuk , A comparison of the results of multistate classification of economic objects using discriminant analysis and artificial neural networks.....	274
Jerzy Krawczuk , Effectiveness of classification methods in S&P500 stock index direction changes forecasting.....	282
Anna Czapkiewicz, Beata Basiura , The simulation study of the utility of the Copula-GARCH models for clustering financial time series.....	290
Radosław Pietrzyk , Timing and selectivity in mutual funds performance measurement.....	305
Aleksandra Witkowska, Marek Witkowski , Use of the Panzar-Rosse method to assess of the competition level in the cooperative banks sector.....	314
Marcin Pelka , Ensemble learning with the application of <i>boosting</i> in symbolic data analysis.....	322
Justyna Wilk , Comparative study of symbolic data classification software.....	332
Tomasz Bartłomowicz, Justyna Wilk , Application of symbolic data analysis methods for domain database searching.....	341
Kamila Migdał-Najman , A proposal of hybrid clustering method based on self-learning networks.....	351
Dorota Rozmus , Comparison of accuracy of spectral clustering and cluster ensembles stability based on bagging idea.....	360
Krzysztof Najman , A dynamic grouping based on self-learning GNG networks.....	369
Małgorzata Misztal , Influence of data imputation methods on the results of object classification using classification trees in the case of small data sets – simulation assessment.....	379
Mariusz Kubus , The application of pre-conditioning of explanatory variable for feature selection.....	386
Barbara Batóg, Jacek Batóg , Application of discriminant analysis to the identification of factors determining the rate of return on the capital market.....	395

Katarzyna Wójcik, Janusz Tuchowski , Comparative analysis of text documents similarity measures based on frequency matrix and based on domain knowledge.....	405
Iwona Staniec , Factor analysis in the identification of areas that determine the improvement of management systems in Polish organizations.....	415
Marek Lubicz, Maciej Zięba, Adam Rzechonek, Konrad Pawełczyk, Jerzy Kołodziej, Jerzy Błaszczyk , Comparative analysis of selected data mining approaches to the classification of medical data with missing values (covariates).....	425
Iwona Foryś , The log-linear analysis using to select the factors determining the attractiveness of the price of flats on the secondary market on the example of local housing market.....	435
Ewa Genge , Trimming approach to the mixtures of normal distributions.....	443
Jerzy Korzeniewski , Efficiency assessment of Ichino method and mean value method of selecting variables in cluster analysis.....	450
Andrzej Dudek , SMS – proposal of new clustering algorithm.....	459
Artur Mikulec , Evaluation methods for the grouping result in cluster analysis.....	468
Małgorzata Machowska-Szewczyk , Fuzzy clustering algorithm for objects described by symbolic or fuzzy variables.....	478
Artur Zaborski , PROFIT analysis and its using in the research of preferences.....	487
Karolina Bartos , Cluster analysis of selected countries due to the structure of their citizens' consumer expenditures – the use of Kohonen networks.....	495
Barbara Batóg, Magdalena Mojsiewicz, Katarzyna Wawrzyniak , Classification of households according to the impulses of concluding the insurance contract by means of qualitative variable models.....	504
Izabela Kurzawa , The application of LA/AIDS model to examine price elasticities of demand of households in the urban-rural relationship.....	512
Aleksandra Luczak, Feliks Wysocki , Linear ordering methods of objects described by a set of metric and ordinal characteristics.....	522
Agnieszka Sompolska-Rzechuła , The comparison of the classical and positional taxonomic analysis of the quality of life differentiation in Zachodniopomorskie voivodeship.....	531
Joanna Banaś, Małgorzata Machowska-Szewczyk , Evaluation of intensity of mailboxes using with the ordered probit model.....	540
Iwona Bąk , Segmentation of pensioners and annuitants households in terms of expenditures on recreation and culture.....	551
Aneta Becker , Application of ANP method to organize Polish voivodships in terms of dynamics of the use of ICT in 2008-2010.....	561
Katarzyna Dębowska , The classification of sectors' financial situation using the methods of multivariate statistical analysis.....	570

Anna Domagała , Proposal of a new method for variable selection in DEA models (combined forward stepwise selection method).....	579
Henryk Gierszal, Karina Pawlina, Maria Urbańska , Statistical analysis in demand research of ICT services in mobile networks.....	589
Hanna Gruchociak , Construction of regression estimator for two-level data	600
Tomasz Klimanek, Marcin Szymkowiak , Application of spatial models in indirect estimation of some labor market characteristics	609
Jarosław Lira , Forecasting of hog livestock production profitability in Poland	618
Christian Lis , The utilization of taxonomic methods in the appraisal of competitiveness of south Baltic ports	627
Beata Bieszk-Stolorz, Iwona Markowicz , The application of the multinomial logit model in evaluating employment odds for the unemployed job seekers	636
Lucyna Przezbórska-Skobiej, Jarosław Lira , Agritourism space of Poland and its valuation.....	645
Paweł Ulman , Model of expenses distribution and demand functions.....	654
Maria Urbańska, Tadeusz Mizera, Henryk Gierszal , Methods of statistical analysis in research of molluscs	663

Henryk Gierszal

Uniwersytet im. Adama Mickiewicza w Poznaniu

Karina Pawlina

Instytut Technik Telekomunikacyjnych i Informatycznych

Maria Urbańska

Uniwersytet Przyrodniczy w Poznaniu

ANALIZA STATYSTYCZNA W BADANIACH ZAPOTRZEBOWANIA NA USŁUGI TELEINFORMATYCZNE SIECI ŁĄCZNOŚCI RUCHOMEJ

Streszczenie: W artykule zaprezentowano możliwości wykorzystania statystycznych analiz wielu zmiennych w badaniach popytu na usługi telekomunikacyjne na przykładzie planowanej ogólnokrajowej radiowej sieci łączności dyspozytorskiej. Wyniki przedstawiono dla analizy kanonicznej, dyskryminacyjnej, czynnikowej, korespondencji oraz skupień, które pozwoliły zidentyfikować zachowania użytkowników determinujące zapotrzebowanie na tego rodzaju usługi.

Słowa kluczowe: badania popytu, rynek usług łączności ruchomej, wielowymiarowa analiza danych.

1. Wstęp

Przy badaniach rynku usług telekomunikacyjnych w zakresie popytu, zachowań, preferencji czy satysfakcji użytkowników powszechnie stosuje się analizę statystyczną, dzięki której możliwe jest uchwycenie relacji, tendencji, potrzeb oraz niedociągnięć natury podażowej, technicznej czy organizacyjnej. Ponieważ w kraju ma zostać wkrótce wybudowana radiowa sieć łączności dyspozytorskiej (trankingowej) o najwyższym możliwym w służbach cywilnych stopniu pewności działania i pewności zasięgu, postanowiono zbadać popyt na usługi telekomunikacyjne, które może ona zaoferować. Dzięki temu operator przyszłej sieci będzie mógł lepiej zaplanować harmonogram uruchamiania kolejnych usług. Jednym z głównych celów analizy było zatem określenie, jaka część funkcjonujących na rynku podmiotów byłaby

skłonna zastąpić usługi oferowane przez telefonię komórkową (o ile ich używa już dziś) cyfrową siecią łączności trunkingowej standardu TETRA (*TErrestrial Trunked RAdio*). Dzięki rozwiązaniom technicznym system ten pozwala m.in. na:

- szyfrowaną (lub nieszyfrowaną) komunikację głosową;
- integrację usług z informacją lokalizacyjną pobieraną za pomocą systemu GPS;
- dostęp do baz danych (przez Internet czy instytucjonalne intranety z przepływnością do 28,8 kbit/s);
- przesyłanie komunikatów statutowych (odpowiednik SMS w standardzie GSM);
- implementację WAP.

Wraz z budową sieci trunkingowej powstanie możliwość świadczenia podmiotom publicznym i prywatnym usług głosowych oraz usług transmisji danych. Potencjał takiej oferty komercyjnej będzie większy, o ile uda się wybudować homogeniczną sieć standardu TETRA, tzn. opartą na systemie jednego dostawcy lub chociaż z pełną możliwością realizacji usługi *roaming* krajowy. Jeśli powstałyby odrębne sieci wyspowe oparte na systemach wielu producentów, wówczas trzeba będzie zapewnić możliwość interoperacyjności, co będzie stanowiło dodatkowy koszt w tym przedsięwzięciu.

2. Metoda badawcza

Analiza popytu rynku na usługi trunkingowe została przygotowana na podstawie badań ankietowych, które przeprowadzono na grupie przedsiębiorstw podzielonej na 15 typów w zależności od wielkości podmiotu i segmentu rynku¹. Wśród badanych podmiotów wyróżniono trzy grupy pod względem ich wielkości wyrażonej liczbą osób zatrudnionych zgodnie z przyjętą przez Wspólnotę Europejską klasyfikacją przedsiębiorstw przedstawioną w [WE 2004]. W badaniu pominięto mikroprzedsiębiorstwa zatrudniające do 9 pracowników.

Ponadto wyróżniono pięć segmentów rynku, takich jak: instytucje, usługi, handel, produkcja i segment technologii komunikacyjno-informatycznych (TKI), które wyodrębniono według Polskiej Klasyfikacji Działalności [PKD 2007] zgodnie z tab. 3. Segment TKI obejmował wszelką działalność (usługi, handel, produkcja) związaną z telekomunikacją i informatyką. Segment został potraktowany oddzielnie jako specyficzny pod względem wykorzystania usług telekomunikacyjnych oraz jako posiadający szczególne potrzeby. Wszystkie grupy są rozłączne, tzn. w sektorach: instytucje, usługi, handel i produkcja nie ma podmiotów z branży TKI.

W ramach badania zebrano 277 ankiet. Liczba obserwacji poszczególnych grup wielkościowych i segmentów rynku została przedstawiona w tab. 1. Przyjęte minimalne liczby obserwacji zostały ustalone w taki sposób, by odzwierciedlać strukturę rynku (podmiotów większych jest mniej) oraz by jednocześnie każda z grup była wystarczająco liczna, aby wyniki badań odzwierciedlały stan rynku.

¹ Prace były realizowane w ramach projektu NCBiR KB/125/13193/IT1-B/U/08.

Tabela 1. Liczba przedsiębiorstw biorących udział w badaniu w zależności od wielkości i sektora

	Małe (od 10 do 49)	Średnie (od 50 do 249)	Duże (250 i więcej)	Ogółem w sektorze
Instytucje	20	15	15	50
Usługi	32	22	11	65
Handel	27	18	10	55
Produkcja	26	17	10	53
TKI	22	16	16	54
OGÓLEM	127	88	62	277

Źródło: opracowanie własne.

W badaniach wyodrębniono 34 zmienne, które zestawiono w tab. 2. Do oceny zapotrzebowania na usługi sieci dyspozytorskiej wykorzystano metody statystycznej analizy wielowymiarowej obejmującej analizę kanoniczną, dyskryminacji, czynnikową, korespondencji i skupień, co przedstawiono w pkt 3. Do przeprowadzenia tych badań wykorzystano pakiet statystyczny Statistica.

Tabela 2. Zmienne opisujące ankietowane przedsiębiorstwa

Lp.	Zmienna	Opis	Lp.	Zmienna	Opis
1	2	3	4	5	6
1	MIASTO_W	czy siedziba firmy jest w mieście wojewódzkim	2	WIELKOŚĆ	wielkość firmy
3	SEKTOR	segment rynku według PKD	4	TYP	typ przedsiębiorstwa
5	LICZBA_P	liczba pracowników w firmie	6	KOMÓRKI	czy używa się telefonów komórkowych
7	GŁOS	intensywność korzystania z usług głosowych	8	SMS	intensywność korzystania z wiadomości SMS
9	MMS	intensywność korzystania z wiadomości MMS	10	INTENSYW	intensywność korzystania z usług głosowych i wiadomości
11	INTERNET	intensywność korzystania z Internetu	12	TRANSMIS	intensywność korzystania z transmisji danych
13	INTENSYW_DANE	intensywność korzystania z usług danych	14	ŁĄCZNA_I	sumaryczna intensywność
15	TEL	liczba telefonów komórkowych	16	TEL_3_LA	prognoza liczby telefonów komórkowych za 3 lata
17	TEL_5_LA	prognoza liczby telefonów komórkowych za 5 lat	18	MOD	liczba modemów dostępu do Internetu w sieci telefonii komórkowej

1	2	3	4	5	6
19	MOD_3_LA	prognoza liczby modemów do Internetu za 3 lata	20	MOD_5_LA	prognoza liczby modemów do Internetu za 5 lat
21	M2M	liczba modemów M2M w sieci telefonii komórkowej	22	M2M_3_LA	prognoza liczby modemów M2M za 3 lata
23	M2M_5_LA	prognoza liczby modemów M2M za 5 lat	24	CZY_TETRA	czy firma chciałaby korzystać z sieci TETRA
25	POTRZE_T	potrzebna jest transmisja szerokopasmowa	26	ZASIĘG	potrzebna jest sieć o dużym zasięgu
27	JAKOŚĆ	potrzebna jest sieć o wysokiej jakości działania	28	CENA	koszt usług nie może być wysoki
29	JEST_OK	obecny dostawca usług jest satysfakcjonujący	30	BRAK_POT	brak potrzeb do zmiany sieci
31	BRAK_OFE	brak ofert od dostawców alternatywnych	32	NIE_LUBI	firma nie lubi zmian
33	ZOBOWIĄZ	firma ma długoterminowe zobowiązania wobec obecnego dostawcy	34	ROZM_ZAG	prowadzi się dużo rozmów zagranicznych

M2M (*Machine to Machine*) — modem radiowy pozwalający na transmisję danych.

Źródło: opracowanie własne.

3. Wyniki

3.1. Analiza kanoniczna

Punktem wyjścia w analizie kanonicznej (uogólnienie regresji wielokrotnej) [Galanc i in. 2011] są pary zmiennych kanonicznych o określonym stopniu skorelowania. Zmienne: ‘miasto wojewódzkie’, ‘wielkość przedsiębiorstwa’, ‘sektor działalności przedsiębiorstwa’, ‘wykorzystywanie telefonii komórkowej’, ‘intensywność korzystania z usług głosowych i wiadomości’ oraz ‘intensywność korzystania z usług danych’ są zmiennymi prognozującymi (**jeden zbiór**); szukamy powiązania z ‘gotowością korzystania z łączności trunkingowej TETRA’ (drugi zbiór). Wartość testu χ^2 sprawdzającego istotność tej jedynej korelacji kanonicznej wynosi 0,336, poziom prawdopodobieństwa zaś $p = 0,016$. Otrzymujemy jeden pierwiastek kanoniczny, gdyż tylko jedna zmienna została przypisana do jednego ze zbiorów. Wagi kanoniczne dla tej jednej zmiennej kanonicznej zebrano w tab. 4.

Tabela 3. Kwalifikacja segmentów rynku według głównych sekcji PKD 2007

Segment rynku	Rodzaj działalności według głównych sekcji PKD 2007
Instytucje	O: administracja publiczna i obrona narodowa; obowiązkowe zabezpieczenia społeczne; P: edukacja; Q: opieka zdrowotna i pomoc społeczna;
Usługi	I: działalność związana z zakwaterowaniem i usługami gastronomicznymi; J.58-60: informacja i komunikacja; H: transport i gospodarka magazynowa; K: działalność finansowa i ubezpieczeniowa; L: działalność związana z obsługą rynku nieruchomości; M: działalność profesjonalna, naukowa, i techniczna; N: działalność w zakresie usług administrowania i działalność wspierająca; R: działalność związana z kulturą, rozrywką i rekreacją; S: pozostała działalność usługowa; T.98.2: Gospodarstwa domowe świadczące usługi na własne potrzeby; U: organizacje i zespoły eksterytorialne;
Handel	G: handel hurtowy i detaliczny; naprawy pojazdów mechanicznych, włączając motocykle; T.97: gospodarstwa domowe zatrudniające pracowników;
Produkcja	A: rolnictwo, leśnictwo łowiectwo i rybactwo; B: górnictwo i kopalnictwo; C: przetwórstwo przemysłowe; D: wytwarzanie i zaopatrywanie w energię elektryczną, gaz, parę wodną, gorącą wodę i powietrze do układów klimatyzacyjnych; E: dostawa wody; gospodarowanie ściekami i odpadami oraz działalność związana z rekultywacją; F: budownictwo; T.98.1: gospodarstwa domowe produkujące wyroby na własne potrzeby;
TKI	J.61-63: informacja i komunikacja — wszelka działalność (handel, produkcja, usługi), która związana jest z telekomunikacją i informatyką

Źródło: opracowanie własne.

Tabela 4. Wagi kanoniczne zmiennych prognozujących

Zmienna prognozująca	Waga kanoniczna
MIASTO_W	0,185295
WIELKOŚĆ	-0,636183
SEKTOR	-0,356206
KOMÓRKI	-0,011670
INTENSYW	-0,720763
INTENSYW_DANE	0,202740

Źródło: opracowanie własne.

Największe bezwzględne wartości wag mają zmienne INTENSYW oraz WIELKOŚĆ. Można zatem przypuszczać, że korelacja między tymi dwiema zmiennymi wpłynęła na powstanie korelacji kanonicznej pomiędzy obydwoma zbiorami danych.

Obliczając ładunki czynnikowe (tzn. korelacja między zmiennymi kanonicznymi a zmiennymi w każdym zbiorze – im większy ładunek czynnikowy co do wartości bezwzględnej, tym ważniejsza ta zmienna w zmiennej kanonicznej), przekonujemy się, że duże wartości bezwzględne przyjmuje zmienna WIELKOŚĆ, INTENSYW i SEKTOR – tab. 5.

Tabela 5. Wagi kanoniczne dla zmiennych prognozujących

Zmienna prognozująca	Ładunki czynnikowe
MIASTO_W	0,185295
WIELKOŚĆ	-0,589845
SEKTOR	-0,514547
KOMÓRKI	0,231088
INTENSYW	-0,536289
INTENSYW_DANE	0,111135

Źródło: opracowanie własne.

Korzystając z tych danych, można stwierdzić, że wyznaczona zmienna kanoniczna wyodrębnia 17% wariancji w zbiorze zmiennych prognozujących. Przy tej zmiennej kanonicznej można wyjaśnić tylko 1,9% wariancji tych 6 zmiennych, bo tyle wynosi wartość redundancji.

Jeśli do zbioru zmiennych prognozujących dołączyć także trzy zmienne opisujące prognozy zapotrzebowania na ‘usługę głosową w telefonii komórkowej’, ‘usługę dostępu do Internetu’ oraz ‘usługę transmisji danych’, to ani wartość wariancji wyodrębnionej, ani redundancja nie zwiększyła się.

3.2. Analiza dyskryminacji

Do analizy dyskryminacji – stosowanej w m.in. na potrzeby marketingu [Stanimir 2008] czy profilowania segmentów [Jefmański 2006] – pozwalającej rozstrzygnąć, które zmienne wyróżniają grupę osób chętnych i niechętnych do nabycia usług telefonicznej, posłużono się 5 zmiennymi w modelu: WIELKOŚĆ, INTENSYW, INTENSYW_DANE, MOD_3_LA i M2M_3_LA. Możliwości dyskryminacyjne tego modelu są wysoce istotne ($p < 0,0026$). Wartość cząstkowej lambdy Wilksa wskazuje, że zmienna M2M_3_LA ma największy wkład do ogólnej dyskryminacji; drugą jest INTENSYW_DANE.

3.3. Analiza czynnikowa

Posłużmy się także analizą czynnikową [Sochańska-Kawiecka, Kołakowska 2008], aby zredukować liczbę 28 zmiennych, które opisują populację przedsiębiorstw. Ograniczając liczbę czynników głównych do 11, otrzymano dla nich wartości zebrane w tab. 6. Wartość własna określa wariancję nowych kolejno wyodrębnionych czynników. Wartości te wyrażono także w procentach całkowitej wariancji. Czynniki 1 wyjaśnia niemal 16% wariancji, czynnik 2 — 10% itd.

Tabela 6. Wartości własne i odsetek całkowitej wariancji dla wyznaczonych czynników głównych

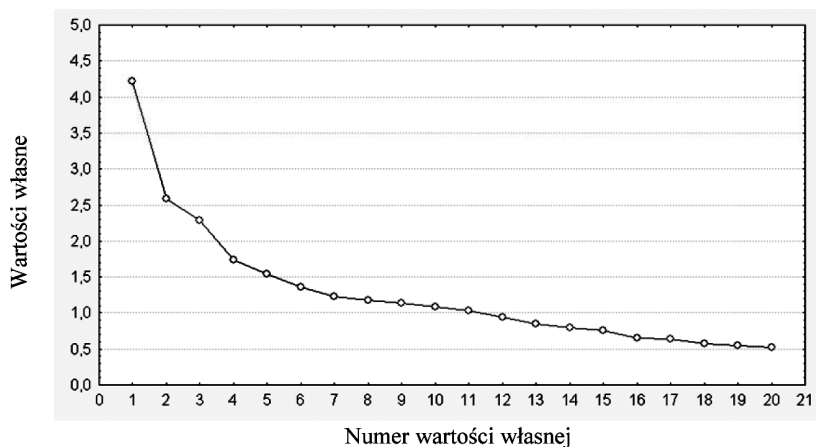
Czynnik główny	Wartość własna	Odsetek całkowitej wariancji
1.	4,24141	15,1479
2.	3,07781	10,9922
3.	2,30947	8,24812
4.	1,89162	6,75579
5.	1,65712	5,91828
6.	1,3619	4,86391
7.	1,23086	4,39591
8.	1,19043	4,25155
9.	1,1387	4,06679
10.	1,09765	3,92017
11.	1,03684	3,70301

Źródło: opracowanie własne.

Liczba 11 czynników wynika z kryterium Kaisera, które jest stosowane najczęściej. Podobną obserwację można zrealizować, robiąc test osypiska zaproponowany przez Cattella (rys. 1). Zwykle test osypiska pozostawia zbyt mało zmiennych w modelu. Widoczne jest to także tutaj, gdyż łagodniejsze opadanie krzywej zaczyna się już przy 7, 8 czynniku.

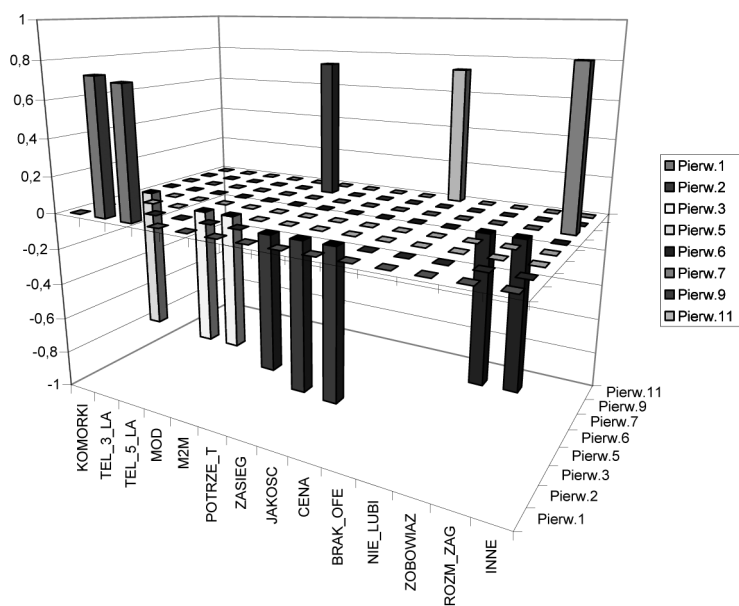
Wykonując analizę składników głównych, określa się korelację między zmiennymi i czynnikami (korelacje te określa się nazwą ładunków czynnikowych). Czynniki 1 jest skorelowany najbardziej z zapotrzebowaniem na modemy do dostępu do Internetu (i transmisji danych). Z kolei czynnik 2 wiąże się m.in. z zapotrzebowaniem na łączność trunkingową.

Wprowadzając rotację struktury czynnikowej [Kossowki 2003], można doprowadzić do bardziej przejrzystego układu ładunków określających poszczególne zmienne. Wybierając strategię rotacji *quartimax* znormalizowaną, otrzymujemy bardziej przejrzysty układ ładunków, w którym czynniki są jasno wyróżnione przez wysokie ładunki dla niektórych zmiennych i niskie ładunki dla innych. Ten model jest bardziej czytelny. Czynniki 1 wyróżnia zapotrzebowanie na usługi telefonii i dostępu do



Rys. 1. Wykres wartości własnych

Źródło: opracowanie własne.



Rys. 2. Pierwotne ładunki czynnikowe o wartości bezwzględnej większej od 0,7

Źródło: opracowanie własne.

Internetu; czynnik 2 – powody, dla których przedsiębiorstwo nie chce skorzystać z usług sieci trunkingowej (są one następujące: nie ma gwarancji porównywalnego zasięgu sieci; jakość usług może nie odpowiadać oczekiwaniom; ceny mogą być

wyższe niż obecna stawki rynkowe); czynnik 3 – obecny stan posiadania modemów do dostępu do Internetu i transmisji M2M; zaś czynnik 4 – chęć posiadania usług sieci trankingowej, z którą wiąże się także zmienna określająca zadowolenie ze współpracy z dotychczasowym dostawcą usług radiokomunikacji ruchomej.

Zamiast obliczać ładunki dla często trudnych do interpretacji czynników ukośnych (nieortogonalnych), można wykorzystać strategię w postaci hierarchicznej analizy czynnikowej. W efekcie takiej analizy okazuje się, że istnieją dwa wtórne czynniki, które wpływają na zmienne. Z czynników głównych (rys. 2) można wyodrębnić obszar: zapotrzebowania na usługi telefonii (czynnik pierwotny 1); powody, dla których przedsiębiorstwo nie chce skorzystać z usług sieci trankingowej (są to następujące powody: nie ma gwarancji porównywalnego zasięgu sieci; jakość usług może nie odpowiadać oczekiwaniom; ceny mogą być wyższe niż obecne stawki rynkowe) (czynnik pierwotny 2); obecny stan posiadania modemów do dostępu do Internetu i transmisji M2M (czynnik pierwotny 3); dodatkowe powody, dla których przedsiębiorstwo nie chce skorzystać z usług sieci trankingowej (są to: istniejące zobowiązania w stosunku do bieżącego dostawcy usług; wykorzystywanie telefonii ruchomej w kontaktach międzynarodowych) (czynnik pierwotny 6).

4. Podsumowanie

Wykorzystane metody analizy statystycznej pozwoliły zidentyfikować najważniejsze czynniki popytowe związane ze świadczeniem usług trankingowych. Redukcja liczby zmiennych pozwoliła wychwycić główne czynniki determinujące chęć zakupu usług w nowej sieci lub przeniesienia się do niej po rezygnacji z umowy z dotychczasowym dostawcą usług.

Zapotrzebowanie na usługi trankingowe silnie zależy od wielkości przedsiębiorstwa, od obecnego stopnia wykorzystywania usług telefonii komórkowej, a także od sektora, w którym funkcjonuje podmiot. Więcej niż co trzecie przedsiębiorstwo w kraju jest skłonne zmienić obecnego dostawcę usług telefonii komórkowej na operatora sieci trankingowej. Największe zainteresowanie zmianą usług telefonii komórkowych na sieć trankingową jest wśród instytucji i w sektorze usług. Niezależnie od wielkości podmiotu najmniejszym zainteresowaniem technologią dyspozytorską charakteryzuje się sektor handlu. Z kolei podmioty sektora TKI, niezależnie od wielkości, charakteryzują się podobnym stopniem zainteresowania sieciami dyspozytorskimi.

Popyt na usługi trankingowe wyznaczają także potrzeby w zakresie usługi dostępu do Internetu i transmisji danych. Przedsiębiorstwa zainteresowane technologią dyspozytorską charakteryzują się wyższym poziomem wykorzystania usług telekomunikacyjnych.

Literatura

- Galanc T., Kołwzan W., Pieronek J., *A quantitative method supporting the management of the teaching process*, „Operations Research and Decision” 2011, no 1.
- Jefmański B., *Analiza dyskryminacyjna w profilowaniu segmentów rynkowych*, [w:] *Ilościowe i jakościowe metody badania rynku. Pomiar i jego skuteczność*, Prace Naukowe Akademii Ekonomicznej w Poznaniu nr 71, Poznań 2006.
- Kossowki T., *Analiza czynnikowa – problem porównania czynników wyodrębnionych różnymi metodami*, [w:] *Problemy interpretacji wyników metod badawczych w geografii społeczno-ekonomicznej i gospodarce przestrzennej*, H. Rogacki (red.), Bogucki Wydawnictwo Naukowe, Poznań 2003.
- PKD 2007, *Polska klasyfikacja działalności*, www.stat.gov.pl/klasyfikacje/pkd_07/pkd_07.htm.
- Sochańska-Kawiecka M., Kołakowska Z., *Bierni czy aktywni. Sytuacja osób powyżej 45 roku życia na rynku pracy województwa warmińsko-mazurskiego*, Raport końcowy z badań zrealizowanych dla Wojewódzkiego Urzędu Pracy w Olsztynie, 2008.
- Stanimir A., *Wykorzystanie analizy korespondencji w badaniach marketingowych*, Zastosowania metod statystycznych w badaniach naukowych III, 2008.
- Stanisz A., *Podstawy statystyki dla prowadzących badania naukowe*, „Medycyna Praktyczna”, www.mp.pl, 2002.
- WE 2004, Rozporządzenie Komisji (WE) nr 364/2004 z dnia 25 lutego 2004 roku zmieniające Rozp. Komisji (WE) nr 70/2001 – Załącznik I stanowiący wyciąg Zalecenia KE 2003/361/WE z dn. 6.05.2003 dotyczącego definicji mikro-, małych, średnich przedsiębiorstw, OJ L 63/2004, 22.

STATISTICAL ANALYSIS IN DEMAND RESEARCH OF ICT SERVICES IN MOBILE NETWORKS

Summary: In this paper the statistical multi variable analyses are presented that were applied in demand research for telecommunication services on the example of nation-wide radio dispatching network functioning in trunking technology. The results include canonical, discrimination, factor, correspondence and cluster analyses that allowed the identification of user's behavior determining the demand for this type of services.

Keywords: demand research, market of mobile radio services, multivariate data analysis.