

**Małgorzata Łatuszyńska\*****Agata Wawrzyniak\*\*****Barbara Wąsikowska\*\*\***

Uniwersytet Szczeciński

**Erika Galindo\*\*\*\*****Judith Cruz Sandoval\*\*\*\*\***

Universidad de las Américas Puebla

**TEORIA ZBIORÓW PRZYBLIŻONYCH W WYKRYWANIU  
REGUŁ ZACHOWAŃ ZAKUPOWYCH KOBIECI I MĘŻCZYŹN  
PODCZAS KUPOWANIA TELEFONÓW KOMÓRKOWYCH****Streszczenie**

Zaprezentowane w artykule badania skupiają się na analizie danych dotyczących preferencji zakupowych kobiet i mężczyzn. Główny nacisk położono na metodę użytą w badaniu – teorię zbiorów przybliżonych. Metodę tę zastosowano do identyfikacji reguł zachowania kobiet i mężczyzn podczas kupowania telefonów komórkowych i akcesoriów. Otrzymane wyniki pozwalają na sformułowanie wniosku, że teoria zbiorów przybliżonych może być z powodzeniem użyta w praktyce jako skuteczne narzędzie dla tego typu analizy danych. Stworzona baza reguł dotyczących preferencji zakupowych kobiet i mężczyzn może służyć firmom produkującym telefony komórkowe i akcesoria jako źródło wiedzy informujące, na co zwracają uwagę kobiety i mężczyźni przy zakupie oferowanych produktów.

**Słowa kluczowe:** zachowanie konsumentów, preferencje, teoria zbiorów przybliżonych, data mining

---

\* mlat@wneiz.pl

\*\* agataw@wneiz.pl

\*\*\* barbara.wasikowska@wneiz.pl

\*\*\*\* erika.galindo@udlap.mx

\*\*\*\*\* judith.cruz@udlap.mx

## **Wprowadzenie**

Badanie preferencji stanowi obecnie jeden z istotniejszych obszarów zainteresowań naukowców zajmujących się ekonomią i zarządzaniem. Jest to główna kategoria badawcza w modelowaniu zachowania konsumenta na rynku, w którym bierze się pod uwagę racjonalność podejmowania decyzji. Terminu „preferencje” używa się powszechnie w literaturze naukowej, jednak jego znaczenie różni się w zależności od danej dyscypliny. Preferencje są podstawowym terminem w teorii ekonomii, szczególnie w teorii wyboru konsumenta. Zazwyczaj przyjmuje się, że preferencje odzwierciedlają i formalizują osobiste upodobania konsumenta i nie zależą od ekonomicznych czynników (takich jak ceny dóbr, budżet konsumenta), lecz wyłącznie od subiektywnych odczuć (takich jak: zadowolenie, satysfakcja, szczęście) albo użyteczności, jaką mu zapewniają. Indywidualne preferencje pozwalają konsumentowi dokonywać wyborów w sytuacji istnienia wielu dostępnych wariantów. Stanowią one swego rodzaju skalę ocen. Warto w tym miejscu zauważyć, że konsument nie ma stałej skali preferencji, zatem może zmieniać postawy wobec określonych produktów (Woś, Racočka i Kasperek-Hoppe, 2004, s. 16). Z kolei, z punktu widzenia przedsiębiorstwa, poznanie preferencji obecnych i potencjalnych klientów, stylu życia oraz zwyczajów zakupowych pozwala na przygotowanie i zrealizowanie skutecznego programu oddziaływania na nabywców – strategii marketingowej.

W marketingu badanie preferencji konsumentów prowadzi się zarówno na podstawie obserwacji historycznych, jak również danych o charakterze antycypacyjnym, opisujących intencje konsumentów. Te dwa źródła informacji o preferencjach pozwalają na wyróżnienie ich dwóch rodzajów: ujawnionych oraz wyrażonych. Preferencje ujawnione prezentują rzeczywiste decyzje rynkowe konsumentów. Do analizy tego typu preferencji wykorzystuje się przede wszystkim dane historyczne, np. dane z przeprowadzonych *a posteriori* sondaży dotyczących wyborów rynkowych wcześniej dokonanych przez klientów. Wykorzystywany jest również materiał statystyczny zgromadzony w wyniku rejestracji rzeczywistych decyzji zakupowych. Preferencje wyrażone odnoszą się do deklarowanych (hipotetycznych) zachowań rynkowych konsumentów. W tym przypadku głównym źródłem informacji są dane zgromadzone *a priori* za pomocą sondaży (Szymańska, 2011, s. 1–2).

Celem artykułu jest prezentacja możliwości analizy danych marketingowych z zastosowaniem jednej z metod sztucznej inteligencji, jaką jest teoria

zbiorów przybliżonych. Dane niezbędne do przeprowadzenia badania zostały zebrane przy użyciu metody ankietowej w listopadzie 2013 roku w Meksyku i dotyczyły wyrażonych preferencji zakupowych kobiet i mężczyzn. W artykule weryfikuje się pogląd, że zastosowanie teorii zbiorów przybliżonych do analizy danych marketingowych może stanowić znaczne uzupełnienie w stosunku do badań poznawczych przeprowadzanych z wykorzystywaniem tradycyjnych metod.

### 1. Zbiory przybliżone w definiowaniu reguł zachowania konsumentów

Teoria zbiorów przybliżonych, zapoczątkowana przez Zdzisława Pawlaka w latach osiemdziesiątych ubiegłego wieku i rozwijana przez wiele lat zarówno przez swojego twórcę, jak i innych badaczy, stanowi teoretyczną podstawę różnych metod i narzędzi informatycznych, umożliwiających m.in. indukcję reguł decyzyjnych czy też redukcję zbiorów danych. Teoria ta zyskała dużą popularność i jest obecnie z powodzeniem stosowana w eksploracji danych i odkrywaniu wiedzy, złożonych zadaniach klasyfikacji oraz w komputerowych systemach wspomagania decyzji. Dziedziny, w których dotychczas znalazła zastosowanie, to m.in. medycyna, farmakologia, finanse i marketing (w tym przede wszystkim badanie rynku), rozpoznawanie mowy i obrazów, lingwistyka, sztuczna inteligencja.

Na popularność tej metody złożyło się wiele czynników. Przede wszystkim teoria zbiorów przybliżonych wraz z wykorzystującymi ją algorytmami pozwala analizować duże zbiory danych, umożliwiając przy tym redukcję zbioru cech opisujących obiekty do niezbędnego minimum (tzw. redukt zbioru atrybutów), usuwając niespójność w danych, jeśli coś takiego nastąpiło, oraz generując z danych tzw. reguły minimalne. Można zatem powiedzieć, że pozwala na odzwierciedlenie postępowania eksperta z danej dziedziny, który zazwyczaj swą wiedzę potrafiłby najszybciej przedstawiać jako takie właśnie reguły (w postaci zapisów: *if ... and ... then...*). Za zalety teorii zbiorów przybliżonych należy uznać to, że:

- a) nie wymaga założeń odnośnie do danych (np. prawdopodobieństwa czy rozmytości);
- b) zawiera szybkie algorytmy analizy danych;
- c) ułatwia interpretację wyników;
- d) charakteryzuje się znaczną prostotą matematyczną (Pawlak, 2004, s. 2).

## 2. Zastosowana procedura badawcza i wyniki badań

W niniejszym artykule przedstawiono sposób użycia teorii zbiorów przybliżonych do analizy preferencji kobiet i mężczyzn związanych z zakupem telefonów komórkowych oraz do zbudowania na jej podstawie bazy reguł zachowań. Dane empiryczne zostały zebrane metodą ankietową wśród klientów domu towarowego Palacio de Hierro w mieście Puebla w Meksyku. Ankieta zawierała pięćdziesiąt pytań zamkniętych, do których odpowiedzi opracowano z zastosowaniem pięciostopniowej skali Likerta. Kwestionariusz był podzielony na osiem sekcji, w których pytania dotyczyły odpowiednio: ogólnych cech produktu, ceny, promocji, strategii dystrybucji (czyli klasyczny „marketing mix”, inaczej „4p”: *product, price, place, promotion*), uwarunkowań społecznych, czynników psychologicznych, procesu decyzyjnego i satysfakcji klienta. Pytania o dane demograficzne umieszczono na końcu kwestionariusza, aby zapewnić lepszy wskaźnik zwrotności (wielu ludzi nie chce bowiem ujawniać swoich zarobków ze względów bezpieczeństwa i widząc takie pytanie na początku ankiety, może bać się wzięcia udziału w badaniu). Z respondentami kontaktowano się bezpośrednio w sklepie w listopadzie 2013 roku. W sumie udało się zgromadzić dwieście pięćdziesiąt siedem odpowiedzi dotyczących mobilnych rozwiązań.

Uzyskane dane zostały ułożone w postaci tzw. pierwotnej tablicy informacyjnej (tabela 1). Przez tablicę informacyjną rozumiemy uporządkowaną piątkę  $T = (U, Q, D, V, f)$ , gdzie  $D$  jest zbiorem cech (atrybutów) decyzyjnych,  $U$  – zbiorem obiektów (przykładów) zwanym *uniwersum*,  $Q$  – zbiorem cech (atrybutów warunkowych),  $V = Uq \in Q$  i  $Vq$  – to zbiory wszystkich możliwych wartości cech, natomiast  $f: U \times Q \rightarrow V$  – jest funkcją informacyjną (Rutkowski, 2005, s. 21). Tablicowy sposób reprezentacji danych ma dwie zasadnicze cechy: uniwersalność (pozwala na gromadzenie i przechowywanie różnorodnych danych) oraz efektywność (umożliwia w łatwy sposób komputerową analizę tak zapisanych danych).

W pierwszej kolumnie tabeli 1 zawarto numery nadane poszczególnym osobom biorącym udział w ankiecie (od p1 do p257). W następnych czterdziestu dziewięciu kolumnach znajdują się wartości atrybutów warunkowych. Atrybuty warunkowe to pytania zawarte w ankiecie, natomiast wartości tych atrybutów – odpowiedzi, jakich udzielili respondenci.

Tabela 1

## Fragment pierwotnej tablicy informacyjnej

Nr	Płeć	Kupuję telefony komórkowe i akcesoria wysokiej jakości	Kupuję telefony komórkowe i akcesoria ekskluzywnych marek	...	Lokalizacja stoiska z telefonami komórkowymi w sklepie	Wiek
p1	mężczyzna	nie zgadzam się	nie zgadzam się	...	całkowicie usatysfakcjonowany	25–34
p2	mężczyzna	zdecydowanie się zgadzam	zdecydowanie się zgadzam	...	całkowicie usatysfakcjonowany	35–44
p3	kobieta	nie zgadzam się	nie zgadzam się	...	całkowicie usatysfakcjonowany	25–34
p4	mężczyzna	raczej się zgadzam	nie mam zdania	...	usatysfakcjonowany	18–24
p5	mężczyzna	zdecydowanie się zgadzam	raczej się zgadzam	...	usatysfakcjonowany	25–34
p6	mężczyzna	raczej zgadzam się	raczej nie zgadzam się	...	usatysfakcjonowany	25–34
p7	mężczyzna	raczej zgadzam się	raczej się zgadzam	...	usatysfakcjonowany	35–44
p8	mężczyzna	zdecydowanie się zgadzam	raczej się zgadzam	...	całkowicie usatysfakcjonowany	35–44
p9	mężczyzna	nie zgadzam się	raczej się zgadzam	...	obojętny	54 i więcej
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
p255	mężczyzna	zdecydowanie się zgadzam	zdecydowanie się zgadzam	...	całkowicie usatysfakcjonowany	18–24
p256	kobieta	zdecydowanie się zgadzam	raczej się zgadzam	...	całkowicie usatysfakcjonowany	54 i więcej
p257	kobieta	raczej się zgadzam	nie mam zdania	...	obojętny	18–24

Źródło: opracowanie własne.

W ostatniej kolumnie umieszczono atrybut decyzyjny informujący o tym, do jakiego przedziału wiekowego należy osoba badana. W kolejnym kroku, zgodnie z teorią zbiorów przybliżonych, dane zebrane w pierwotnej tablicy informacyjnej

poddano dyskretyzacji oraz kodowaniu. Sposób kodowania wartości poszczególnych atrybutów pokazano w tabeli 2.

Tabela 2

## Dyskretyzacja wartości atrybutów pierwotnej tablicy informacyjnej

Sposób kodowania atrybutu „płeć”	Sposób kodowania atrybutów od q2 do q37	Sposób kodowania atrybutów od q38 do q49	Sposób kodowania atrybutu decyzyjnego (wiek)
Symbol atrybutu: q1 M – (mężczyzna) K – (kobieta)	Symbol atrybutu: q2...q37 a1 – zdecydowanie się nie zgadzam a2 – raczej się nie zgadzam a3 – nie mam zdania a4 – raczej się zgadzam a5 – zdecydowanie się zgadzam	Symbol atrybutu: q38 ... q49 a1 – całkowicie nieusatysfakcjonowany a2 – nieusatysfakcjonowany a3 – obojętny a4 – usatysfakcjonowany a5 – całkowicie usatysfakcjonowany	Symbol atrybutu: d a1 – od 18 do 24 lat a2 – od 25 do 34 lat a3 – od 35 do 44 lat a4 – od 45 do 54 lat a5 – 54 i więcej

Źródło: opracowanie własne.

W wyniku przeprowadzonego kodowania otrzymano wtórną tablicę informacyjną (tabela 3).

Tabela 3

## Fragment wtórnej tablicy informacyjnej

	q1	q2	q3	...	q49	d
p1	M	a2	a2	...	a5	a2
p2	M	a5	a5	...	a5	a3
p3	K	a2	a2	...	a5	a2
p4	M	a4	a3	...	a4	a1
p5	M	a5	a4	...	a4	a2
p6	M	a4	a2	...	a4	a2
p7	M	a4	a4	...	a4	a3
p8	M	a5	a4	...	a5	a3
p9	M	a2	a4	...	a3	a5
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
p255	M	a5	a5	...	a5	a1
p256	K	a5	a4	...	a5	a5
p257	K	a4	a3	...	a3	a1

Źródło: opracowanie własne.

Tak przygotowane dane poddano analizie metodą zbiorów przybliżonych, używając do tego programu DAT (Data Analysis Toolbox) działającego w środowisku MATLAB R2010a. Na podstawie wstępnie przeprowadzonych obliczeń, w ramach których wyznaczono zbiory elementarne (czyli zbiory przykładów zawierających te same wartości atrybutów warunkowych) oraz wyznaczono koncepty decyzyjne (czyli zbiory przykładów zawierających tę samą wartość decyzji), podjęto próbę zredukowania zbiorów atrybutów warunkowych. W wyniku przeprowadzonych badań stwierdzono, że wszystkie atrybuty warunkowe są istotne (żaden z atrybutów nie jest zbędny) i stanowią one rdzeń zbioru. Istotność poszczególnych atrybutów warunkowych została zbadana przez obliczenie znormalizowanego wskaźnika istotności (tabela 4).

Tabela 4

Istotność poszczególnych atrybutów warunkowych

Kobiety	Lp.	Atrybut	Istotność
	1	q9	0,1416
2	q34	0,1327	
3	q17	0,0973	
4	q15	0,0796	
5	q14	0,0531	
6	q16	0,0177	
Mężczyźni	1	q19	0,1181
	2	q15	0,0903
	3	q33	0,0903
	4	q3	0,0625
	5	q1	0,0417
	6	q23	0,0278
	7	q27	0,0139

Gdzie:

**Kobiety:**

- q9 – najwyższa cena zawsze oznacza najwyższą jakość,
- q34 – preferuję kupno telefonu komórkowego i akcesoriów po dokonaniu oceny ryzyka, kosztów i mojej chęci kupna,
- q17 – preferuję kupno telefonu komórkowego i akcesoriów w domu towarowym,
- q15 – specjalne oferty (np. prezenty, rabaty) mają istotny wpływ na moje decyzje zakupowe,
- q14 – mój stosunek do sprzedawcy wpływa na moją decyzję zakupu telefonu komórko-

wego i akcesoriów,  
 q16 – preferuję zakup telefonu komórkowego i akcesoriów w najbliższym sklepie.

**Mężczyźni:**

- q19 – preferuję zakup telefonu komórkowego i akcesoriów w formie on-line,  
 q15 – specjalne oferty (np. prezenty, rabaty) mają istotny wpływ na moje decyzje zakupowe,  
 q33 – preferuję kupno tych telefonów komórkowych i akcesoriów, które kupowane są przez moich przyjaciół,  
 q3 – preferuję kupno telefonów komórkowych i akcesoriów z najnowszymi innowacjami i zgodnych z obowiązującymi trendami,  
 q1 – preferuję telefony i akcesoria wysokiej jakości,  
 q23 – preferuję kupno telefonu komórkowego i akcesoriów na podstawie opinii ekspertów i liderów opinii,  
 q27 – preferuję kupno telefonu i akcesoriów spełniających moje oczekiwania.

Źródło: opracowanie własne.

Analiza tabeli 4 wykazała, że istotność każdego z wybranych atrybutów ma wartość powyżej 0, co potwierdza fakt, iż nie można żadnego z nich usunąć z tablicy decyzyjnej – wszystkie atrybuty są w mniejszym lub większym stopniu istotne. Atrybut jest tym istotniejszy, im wartość istotności danego atrybutu jest bliższa jedności.

Kolejnym krokiem w analizie było dokonanie aproksymacji (przybliżenia) zbioru, czyli wyznaczenie dolnych i górnych przybliżeń, brzegu zbioru oraz obszarów pozytywnych i negatywnych zbioru (dokładne wyjaśnienie tych pojęć oraz odpowiednie wzory można znaleźć między innymi w następujących pracach: Mrózek i Płonka, 1999, s. 9–25; Pawlak, 2004., s. 4–5; Rutkowski, 2005, s. 20–45). Na podstawie otrzymanych wyników obliczono współczynnik jakości przybliżenia gamma, który informuje o tym, w ilu procentach analizowane przykłady oraz atrybuty warunkowe pozwalają wyznaczyć reguły pewne. Jakość przybliżenia zbioru  $F$  wyraża się następującym wzorem:

$$\gamma_{\varnothing}(F) = \frac{\text{card}(\text{Pos}_{\varnothing}(F))}{\text{card}(U)},$$

gdzie w liczniku znajduje się liczebność przykładów zawartych w pozytywnym regionie zbioru, a w mianowniku – liczebność przykładów zawartych w całym uniwersum zbioru. W tym przypadku wartość współczynnika gamma wyniosła



1,00, co oznacza, że wszystkie przykłady zebrane w tablicy informacyjnej generują reguły w 100% pewne. Brak jest reguł sprzecznych. W wyniku przeprowadzonej analizy otrzymano reguły, które następnie uproszczono (przez połączenie ze sobą reguł podobnych). W tabeli 5 przedstawiono reguły poparte największą liczbą przypadków (po pięć reguł dla kobiet i mężczyzn).

Tabela 5

## Reguły poparte największą liczbą przypadków

Numer reguły	Reguły dla kobiet
1	Jeżeli $q9 = a3$ i $q14 = a3$ i $q15 = a4$ lub $a5$ i $q16 = a3$ i $q17 = a3$ lub $a4$ i $q34 = a4$ to wiek = $a1$
2	Jeżeli $q9 = a3$ i $q16 = a3$ lub $a4$ i $q17 = a3$ i $q34 = a4$ to wiek = $a2$
3	Jeżeli $q16 = a3$ lub $a4$ i $q17 = a3$ lub $a4$ i $q34 = a4$ to wiek = $a3$
4	Jeżeli $q14 = a4$ i $q16 = a3$ lub $a4$ i $q34 = a4$ to wiek = $a4$
5	Jeżeli $q14 = a4$ i $q16 = a4$ i $q34 = a4$ to wiek = $a5$
Numer reguły	Reguły dla mężczyzn
1	Jeżeli $q1 = a4$ lub $a5$ i $q3 = a4$ lub $a5$ i $q15 = a4$ i $q19 = a1$ lub $a2$ i $q23 = a3$ lub $a4$ i $q27 = a5$ to wiek = $a1$
2	Jeżeli $q1 = a5$ i $q3 = a5$ i $q15 = a4$ i $q19 = a3$ i $q23 = a4$ lub $a5$ i $q27 = a5$ lub $q33 = a3$ to wiek = $a2$
3	Jeżeli $q1 = a4$ lub $a5$ i $q15 = a4$ i $q23 = a4$ i $q27 = a5$ to wiek = $a3$
4	Jeżeli $q1 = a4$ i $q3 = a4$ i $q15 = a3$ i $q19 = a1$ i $q23 = a5$ i $q27 = a4$ to wiek = $a4$
5	Jeżeli $q3 = a4$ i $q19 = a2$ i $q23 = a4$ i $q27 = a5$ to wiek = $a5$

Źródło: opracowanie własne.

Reguły poparte największą liczbą przypadków zostały w następnym kroku odkodowane. Poniżej przedstawiono przykład odkodowanej reguły:

*Reguła zakodowana:*

*Jeżeli  $q3 = a4$  i  $q19 = a2$  i  $q23 = a4$  i  $q27 = a5$  to wiek =  $a5$*

*Reguła po odkodowaniu:*

*Jeżeli mężczyzną interesuje telefon modny (innovacyjny), a równocześnie nie zgadza się on z tym, że najwyższa cena oznacza zawsze najwyższą jakość i ważna jest dla niego opinia liderów opinii, a najważniejsze jest to, aby telefon spełniał jego oczekiwania – to jest to mężczyzna w wieku 54 lat lub starszy.*

Odkodowanie reguł oznacza zakończenie analizy danych metodą zbiorów przybliżonych.

### **3. Wnioski z przeprowadzonych badań**

Na podstawie otrzymanych reguł i analizy istotności poszczególnych atrybutów warunkowych określono, które atrybuty mają największy wpływ na atrybut decyzyjny:

- a) dla kobiet są to: jakość (związana z ceną), ocena ryzyka, kosztów i chęci dokonania zakupu, możliwość kupna telefonu komórkowego i akcesoriów w domu towarowym, specjalne oferty (np. prezenty, rabaty), relacja ze sprzedawcą (rozumiana jako stosunek do niego), nieduża odległość między domem a sklepem;
- b) dla mężczyzn są to: możliwość kupna telefonu komórkowego i akcesoriów w formie on-line, specjalne oferty (np. prezenty, rabaty), preferencje ich przyjaciół, nowoczesność telefonów komórkowych (najnowsze innowacje i trendy), wysoka jakość, rekomendacje ze strony ekspertów i liderów opinii, fakt, że produkt odpowiada oczekiwaniom.

### **Podsumowanie**

Przeprowadzone badania pokazują różnice w podejściu do kupowania telefonów komórkowych u kobiet i mężczyzn. Na przykład, kobiety twierdzą, że jakość (skorelowana z ceną) jest najważniejszym czynnikiem brany pod uwagę podczas kupna nowego telefonu lub akcesoriów. Ponieważ kobiety są głównymi nabywcami w rodzinie, jest zrozumiałe, że odznaczają się większą wrażliwością cenową niż mężczyźni (Schmitt, 2008, s. 72; Barletta, 2014). Mężczyźni natomiast preferują kupno telefonów komórkowych i akcesoriów w formie on-line. Wykazują także większe zainteresowanie specjalnymi ofertami i cechami telefonów (szczególnie najnowszymi trendami i innowacjami). Są również bardziej podatni na wpływ opinii wyrażanych przez ekspertów lub przyjaciół. Otrzymane wyniki wskazują, że w przypadku dużej liczby respondentów (zarówno kobiet, jak i mężczyzn) promocje (np. prezenty, rabaty) odgrywają ważną rolę w podejmowaniu decyzji zakupowych.

Podsumowując, należy stwierdzić, że zauważalne różnice w zachowaniach zakupowych kobiet i mężczyzn wymagają od firm właściwego zdefiniowania ryn-

ku docelowego i określenia grup odbiorców (np. czy reklama ma być skierowana przede wszystkim do kobiet, czy mężczyzn, a może do przedstawicieli obu płci). Prawidłowe zdefiniowanie rynku pozwala bowiem na bardziej efektywną komunikację z konsumentami.

## Bibliografia

- Barletta M. (2014), *Women Control about 80% of Household Spending: A Look at the Numbers*, The TrendSight Group, <http://www.trendsight.com/content/view/40/204/> (09.05.2014).
- Kraft H., Weber J.M. (2012), *A Look at Gender Differences and Marketing Implications*, „International Journal of Business and Social Science”, t. 3, nr 21, s. 247–253.
- Mrózek A., Płonka L. (1999), *Analiza danych metodą zbiorów przybliżonych. Zastosowania w ekonomii, medycynie i sterowaniu*, Akademicka Oficyna Wydawnicza PLJ, Warszawa.
- Pawlak Z. (1982), *Rough sets*, „International Journal of Information and Computer Science”, t. 341 (11), s. 344–356.
- Pawlak Z. (2004), *Zbiory przybliżone, nowa matematyczna metoda analizy danych*, „Miesięcznik Politechniki Warszawskiej” 5/2004, s. 2.
- Rudnicki L. (2012), *Zachowania konsumentów na rynku*, PWE, Warszawa.
- Rutkowski L. (2005), *Metody i techniki sztucznej inteligencji*, Wydawnictwo Naukowe PWN, Warszawa.
- Schmitt G. (2008), *Hunters and Gatherers*, „Dealernews”, t. 44, nr 8, s. 72.
- Szymańska A.I. (2011), *Badania preferencji konsumentów z wykorzystaniem kompozycyjnej metody badań MDPREF*, PWE, „Marketing i Rynek” 10/2011, s. 23–30.
- Wąsikowska B. (2007), *Application of Rough Sets for Identification of Factors Determining Sold Production of Industry*, „Polish Journal of Environmental Studies”, t. 16, nr 4A, s. 372–375.
- Woś J., Racoeka J., Kasperek-Hoppe M. (2004), *Zachowania konsumentów. Teoria i praktyka*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.

**APPLICATION OF ROUGH SET THEORY TO IDENTIFICATION  
PURCHASE BEHAVIOR RULES OF WOMEN AND MEN  
WHEN BUYING MOBILE PHONES**

**Summary**

The research presented in the article was focused on the data analysis concerning purchase preferences of men and women. The main emphasis is put on the method which was used in the research – the rough set theory. This method was applied to identify rules of male and female behavior while buying cellular phones. The received results allow conclusion that the used method of artificial intelligence i.e. rough set method can be successfully used in practice as an effective tool for this type of data analysis. The created basis of purchase preferences rules for men and women can be used as a base of knowledge for companies producing cellular phones and accessories and can be a direction showing what this two groups of consumers pay attention to while buying products offered by these companies.

*Translated by Agata Wawrzyniak*

**Keywords:** consumer behaviour, preferences, rough set theory, data mining