

Józef DZIECHCIARZ*, Marek WALESIAK**

METODA DETEKCJI INTERAKCJI (AID) W BADANIACH MARKETINGOWYCH***

Podstawową dziedziną badań statystycznej analizy wielowymiarowej są badania marketingowe. Wiele technik statystycznej analizy wielowymiarowej jest praktycznie nieobecnych w badaniach teoretycznych i aplikacjach empirycznych. Dotyczy to analizy dyskryminacyjnej, analizy czynnikowej, skalowania wielowymiarowego, analizy pomiaru łącznego oddziaływania zmiennych (*conjoint measurement, conjoint analysis*), metody detekcji interakcji (*automatic interaction detection – AID*) i analizy współwystępowania (*correspondence analysis*). Niektóre z nich nie mają nawet utrwalonych polskich nazw.

Oferta w zakresie technik statystycznej analizy wielowymiarowej musi być znacznie wzbogacona z przeniesieniem akcentów z metod wykorzystujących zmienne mierzone na silnych skalach pomiaru (przedziałowej i ilorazowej) na metody, które posługują się danymi mierzonymi na skalach porządkowych i nominalnych.

Celem artykułu jest zaprezentowanie istoty metody detekcji interakcji oraz wskazanie jej potencjalnych zastosowań w badaniach marketingowych.

Termin **statystyczna analiza wielowymiarowa SAW** odnosi się do grupy metod statystycznych, za pomocą których jednoczesnej analizie są poddane pomiary na przynajmniej dwóch zmiennych opisujących każdy obiekt badania.

Klasyfikacje metod statystycznej analizy wielowymiarowej zaproponowano m.in. w pracach [7], [10] i [11]. W artykule przyjęto jednak klasyfikację, która powstała na gruncie badań marketingowych (rys. 1).

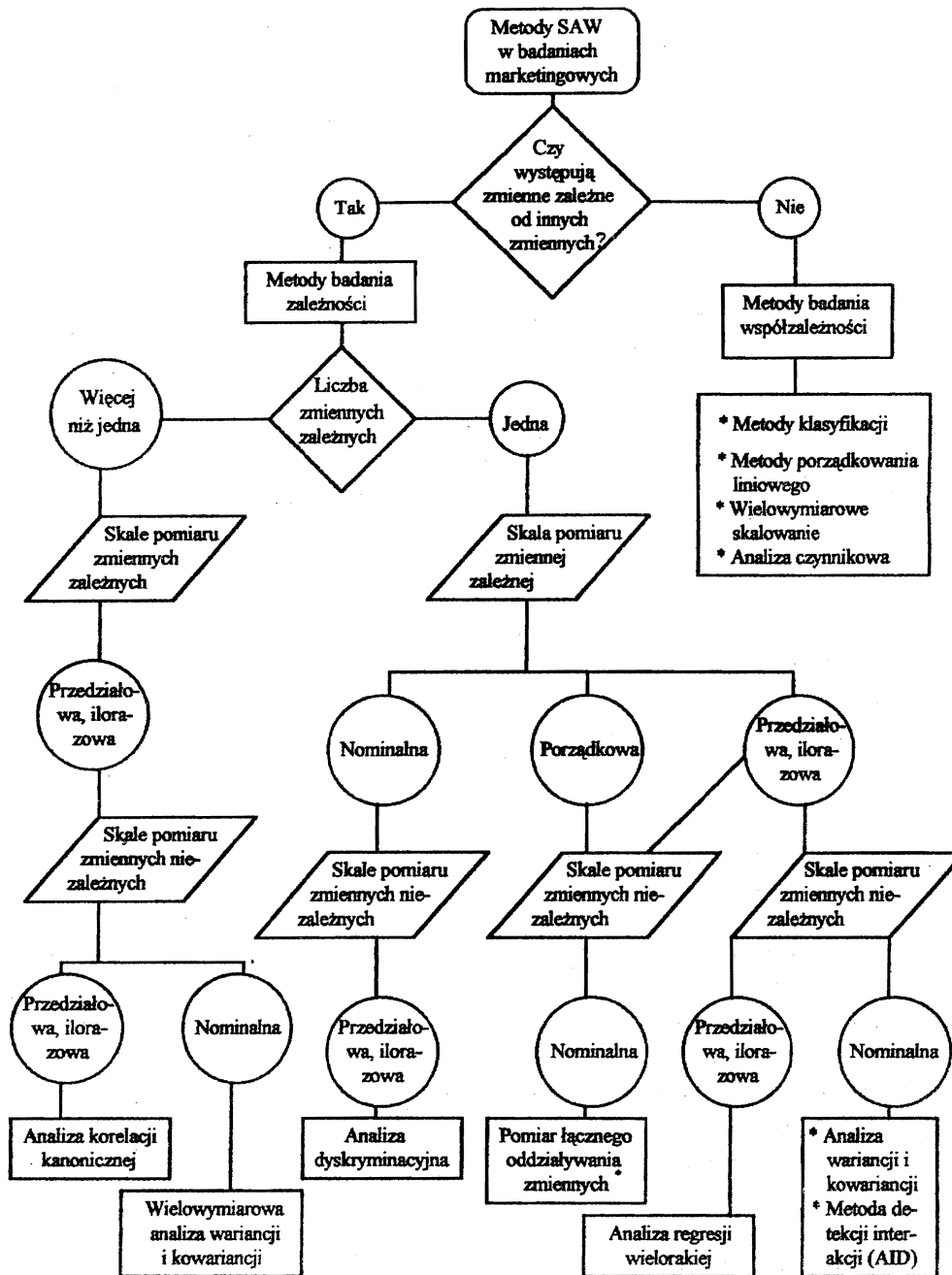
Kinnear i Taylor [12] oraz Green, Tull i Albaum [6, s. 416, 426] przedstawili klasyfikację metod statystycznej analizy wielowymiarowej w badaniach marketingowych według dwóch kryteriów:

1. W zależności od tego, czy w zbiorze badanych zmiennych występuje zmienna zależna (zmienne zależne) od innych zmiennych. Wyróżnia się więc metody badania

* Katedra Ekonometrii Akademia Ekonomiczna im. Oskara Langego, ul. Komandorska 118/120, 53-345 Wrocław.

** Katedra Polityki i Ekonomiki Ochrony Środowiska, Akademia Ekonomiczna im. Oskara Langego, ul. Nowowiejska 3, 58-500 Jelenia Góra.

*** Pracę wykonano w ramach grantu KBN 1-H02B-016-08 pn. *Komputerowo wspomaganie gromadzenie i analiza danych marketingowych*.



* W języku angielskim używa się określeń *conjoint analysis*, *conjoint measurement*.

Rys. 1. Klasyfikacja metod SAW w badaniach marketingowych

Źródło: opracowano na podstawie prac: [6, s. 426]; [12, s. 625]; [13, s. 16]

zależności (*dependence methods*) oraz metody badania współzależności (*interdependence methods*). W pierwszych z nich jedną lub więcej zmiennych określa się jako zmienną (zmiennie) zależną od zbioru pozostałych zmiennych (zwanymi niezależnymi), w drugich zaś, czyli w metodach badania współzależności, nie dzieli się zmiennych na zależne i niezależne. Za pomocą tych metod analizuje się określone wewnętrzne relacje między zmiennymi w zbiorze zmiennych (jeśli przedmiotem badania jest zmienna) lub relacje zachodzące między obiektami opisanymi tymi zmiennymi (gdy przedmiotem badania jest obiekt).

2. W zależności od skali pomiaru zmiennych (skali pomiaru zmiennej lub zmiennych zależnych i skali pomiaru zmiennych niezależnych).

Dane przedstawione w tabeli 1 ilustrują stopień wykorzystania metod statystycznej analizy wielowymiarowej w badaniach marketingowych przedsiębiorstw amerykańskich w 1988 r. Z tabeli wynika, że metoda AID nie należy do metod często stosowanych w badaniach marketingowych. Główną przyczyną tkwi w tym, że wymaga ona dysponowania bardzo dużą próbą statystyczną, obejmującą co najmniej 1000 obserwacji.

Tabela 1

Wykorzystanie metod statystycznej analizy wielowymiarowej w badaniach marketingowych przedsiębiorstw amerykańskich w 1988 r. (w procentach)

Metoda	Wykorzystywana często	Wykorzystywana czasami	Nie wykorzystywana
Analiza regresji	24,28	50,09	25,63
Analiza wariancji i kowariancji	19,51	37,85	42,64
Metody klasyfikacji	8,40	40,63	50,97
Analiza czynnikowa	9,20	31,70	59,10
Pomiar łącznego oddziaływania zmiennych (<i>conjoint analysis</i>)	5,64	30,35	64,01
Analiza dyskryminacyjna	6,86	28,23	64,91
Wielowymiarowe skalowanie	7,45	26,47	66,08
Analiza korelacji kanoniczej	0,98	6,67	92,35
Metoda AID	1,77	4,72	93,50

Źródło: opracowanie własne na podstawie pracy [12].

Metoda detekcji interakcji została zaproponowana w badaniach marketingowych w 1964 r. przez Sonquista i Morgana (por. [8, s. 409]). W metodzie tej jest dana jedna zmienna zależna mierzona na skali przedziałowej lub ilorazowej, której zmienność jest wyjaśniana przez dychotomiczne zmienne niezależne mierzone na skali nominalnej (pierwotnie zmienne mogą być mierzone na skali ilorazowej, przedziałowej, porządkowej i nominalnej, ale na potrzeby tej metody zostają one przekodowane na zmienne dychotomiczne).

Głównym celem tej metody jest podział badanej próby statystycznej na mniejsze, relatywnie jednorodne grupy, mające istotnie mniejszą zmienność niż całkowita próba statystyczna. Miarą zmienności w tej metodzie jest suma kwadratów odchyłeń empirycznych wartości zmiennej zależnej od wartości średniej.

Typowa procedura, wykorzystująca metodę detekcji interakcji w badaniach marketingowych, obejmuje następujące etapy:

1. Określamy wszystkie zmienne niezależne, mające wpływ na wyróżnioną zmienną zależną. Na potrzeby tej metody zmienne mierzone na skali ilorazowej, przedziałowej, porządkowej i nominalnej (wielowariantowe) zostają przekodowane na dychotomiczne (dwuwariantowe) zmienne niezależne.

2. Dla obydwu klas wyróżnionych na podstawie dychotomicznej zmiennej niezależnej obliczamy za każdym razem międzyklasową sumę kwadratów odchyłeń (SS_B) lub wewnątrzklasową sumę kwadratów odchyłeń (SS_W) korzystając w tym celu z analizy wariancji (ANOVA). Jako pierwszą wybieramy tę zmienną binarną, dla której SS_B (SS_W) przyjmuje maksimum (minimum). Jeśli podzielimy wszystkie obserwacje na zmiennej zależnej Y na dwie klasy ($t = 1, 2$ – warianty dychotomicznej zmiennej niezależnej) o liczebnościach n_1 i n_2 , to całkowitą sumę kwadratów odchyłeń wartości zmiennej Y od średniej \bar{Y} w analizie wariancji (ANOVA) rozkłada się następująco:

$$\sum_{i=1}^n \sum_{t=1}^2 (Y_{it} - \bar{Y})^2 = \sum_{t=1}^2 (\bar{Y}_t - \bar{Y})^2 + \sum_{i=1}^n \sum_{t=1}^2 (Y_{it} - \bar{Y}_t)^2.$$

$$SS = SS_B + SS_W,$$

gdzie: SS_B – międzyklasowa suma kwadratów odchyłeń,

SS_W – wewnątrzklasowa suma kwadratów odchyłeń,

SS – całkowita suma kwadratów odchyłeń.

3. Na podstawie wariantów wybranej w kroku 2 zmiennej dychotomicznej dokonujemy podziału obserwacji na zmiennej zależnej na dwie klasy.

4. Powtarzamy krok 2 i 3, aż do wyczerpania wyróżnionych w kroku 1 dychotomicznych zmiennych niezależnych. Każde dwie nowe klasy rozpatrujemy niezależnie. Procedura ta jest kontynuowana do chwili, gdy wystąpi jedna z poniższych sytuacji:

a) otrzymane klasy będą zawierać zbyt mało obserwacji: minimalny próg – to 30 obserwacji;

b) wydzielone klasy będą względnie jednorodne (zmienność wewnątrz klas jest niewielka): dalszy podział danej klasy nie przyniesie więc spodziewanego efektu;

c) nie ma zmiennych niezależnych, które istotnie zwiększają względną jednorodność klas.

Metoda detekcji interakcji (AID) stosowana w badaniach marketingowych służy do (por. np. prace [1], [4], [5, s. 168], [6, s. 539–541], [8], [12]):

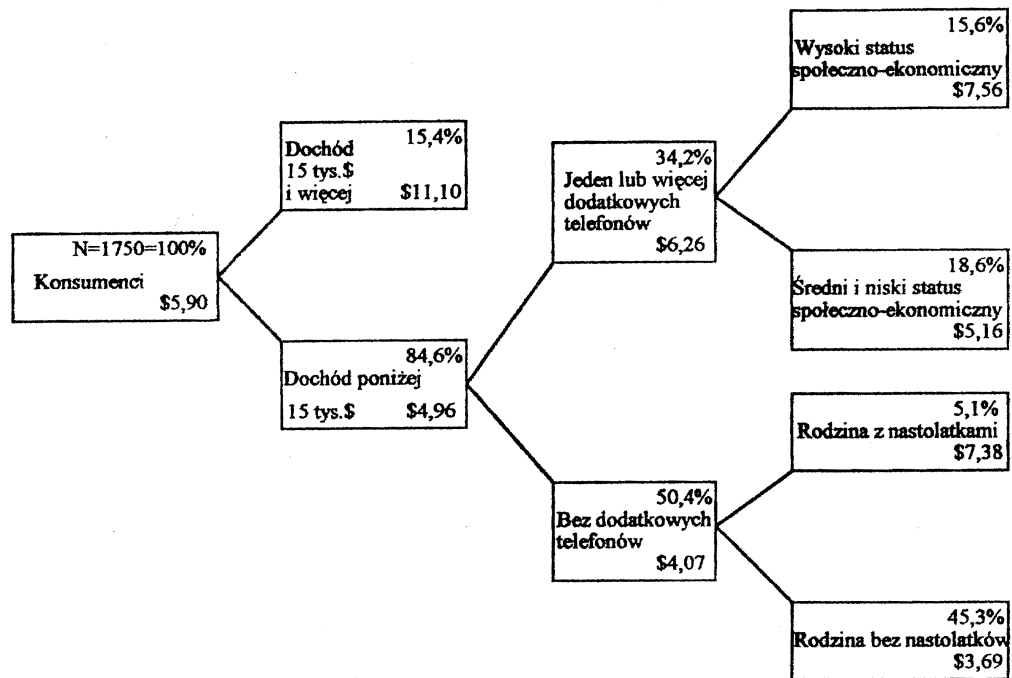
1) określenia relatywnej ważności każdej charakterystyki (zmiennej) w procesie wyboru produktu przez nabywcę,

2) wydzielenia klas (segmentów) potencjalnych konsumentów o zbliżonych preferencjach zakupu.

Assael i Roscoe [2] zastosowali metodę detekcji interakcji w celu identyfikacji segmentów konsumentów usług telefonicznych firmy AT&T. Jako zmienną zależną w badaniu przyjęli przeciętne miesięczne wydatki na międzymiastowe rozmowy telefoniczne w 1972 r. Wśród zmiennych niezależnych, charakteryzujących konsumentów usług telefonicznych, uwzględnili m.in.:

- 1) dochód w 1972 r.;
- 2) wyposażenie mieszkania w dodatkowe telefony;
- 3) status społeczno-ekonomiczny (ustalony na podstawie wykształcenia i wykonywanego zawodu);
- 4) posiadanie dzieci nastoletnich (od 13 do 19 lat);
- 5) typ mieszkania (własne, wynajmowane);
- 6) wielkość rodziny;
- 7) płeć głowy rodziny.

W firmie AT&T zgromadzono dane o 60 tys. konsumentów usług telefonicznych tej firmy. Badaniem objęto próbkę 1 750 osób. Wyniki zastosowania algorytmu metody detekcji interakcji przedstawiono na rys. 2.



Rys. 2. Segmentacja konsumentów usług telefonicznych firmy AT&T

Źródło: praca [2, s. 70].

Tabela 2

Wyniki segmentacji konsumentów usług telefonicznych firmy AT&T uzyskane metodą AID

Nr	Charakterystyka segmentu	Przeciętny rachunek za rozmowy międzymiastowe (w dolarach)	Badana próbka (w pro- centach)	Udział segmentu w sumie wszystkich rachunków (w procentach)
I	Konsumenci o dochodach 15 tys. \$ i więcej	11,10	15,4	29,0
II	Dochód poniżej 15 tys. ; jeden i więcej dodatkowych telefonów; wysoki status społeczno-ekonomiczny	7,56	15,6	20,1
III	Podobnie jak w segmencie II z tym, że konsumenci mają średni i niski status społeczno-ekonomiczny	5,16	18,6	16,2
IV	Dochód poniżej 15 tys. \$; bez dodatkowych telefonów; rodzina z nastolatkami	7,38	5,1	6,4
V	Podobnie jak w segmencie IV z tym, że w rodzinie nie ma nastolatków	3,69	45,3	28,3

Źródło: opracowanie [2, s. 70].

Wyniki segmentacji konsumentów usług telefonicznych firmy AT&T, uzyskane z zastosowaniem algorytmu metody detekcji interakcji, przedstawiono w tab. 2.

Bibliografia

- [1] AAKER D.A., DAY G.S., *Marketing research: private and public sector decisions*, New York, Wiley, 1980.
- [2] ASSAEL H., ROSCOE M.A., *Approaches to market segmentation analysis*, Journal of Marketing, October 1976, vol. 40, 67–76.
- [3] BAKER M.J., *Research for marketing*, London, MacMillan, 1991.
- [4] BEANE T.P., ENNIS D.M., *Market segmentation. A review*, European Journal of Marketing 1987, No. 5, vol. 21, 20–42.
- [5] GARBARSKI L., RUTKOWSKI J., WRZOSEK W., *Marketing*, Warszawa, PWE, 1992.
- [6] GREEN P.E., TULL D.S., ALBAUM G., *Research for marketing decisions*, Englewood Cliffs, Prentice-Hall 1988.
- [7] HAIR J.F., ANDERSON R.E., TATHAM R.L., *Multivariate data analysis with readings*, New York, MacMillan, 1987.
- [8] HOLMES C., *AID comes to the aid of marketing management*, European Journal of Marketing 1980, vol. 14, No 7, 409–413.
- [9] HOOLEY G.J. (ed.), *A guide to the use of quantitative techniques in marketing*, European Journal of Marketing 1980, vol. 14, No. 7.
- [10] JAJUGA K., *Statystyka ekonomicznych zjawisk złożonych – wykrywanie i analiza niejednorodnych rozkładów wielowymiarowych*, Prace Naukowe AE we Wrocławiu nr 371, 1987, seria: Monografie i opracowania nr 39.

- [11] KENDALL M.G., *Multivariate analysis*, London, Griffin, 1975.
- [12] KINNEAR T.C., TAYLOR J.R., *Marketing research. An applied approach*, New York, McGraw-Hill, 1991.
- [13] WALESIAK M., *Statystyczna analiza wielowymiarowa w badaniach marketingowych*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu 1993 nr 654, Seria: Monografie i opracowania nr 101.

Automatic interaction detection in the marketing research

One of the main fields of application of multivariate statistical data analysis is marketing research. Since there was no market in Poland – such applications were absent in the Polish statistical and marketing literature. The situation changes now – many scholars start new projects in that field.

Most of Polish applications of the multivariate statistical data analysis employed regression, clustering and hierarchization methods. Many techniques are completely absent in Polish theoretical research and practical applications. Among them are: discriminant analysis, factor analysis, multidimensional scaling, conjoint analysis, automatic interaction detection, correspondence analysis and others. Some of them even do not have a Polish equivalent in the terminology.

In the meantime there will be demand for marketing research projects. Sample allocation methods, data acquisition and control techniques, standard statistical analysis and multivariate analysis methods have to be in the offer. Those projects have to be compatible with foreign methodology and standards. On the other hand the presentations techniques have to be adequate to the perception ability of the user. Therefore one has to avoid to complex methods of results' presentation. There is the necessity for enrichment of the arsenal of the methods for data processing analysis and results' presentation. Among them the graphical methods have enormous capacity.

The methodological offer of statistical projects has to be enriched with multivariate techniques with transfer of the accents from methods working with metric data to those which are suitable for data from nominal and ordinal scales.

One of such methods is automatic interaction detection. Possible fields of application in the marketing research along with the method itself are presented in the article.

Verified by Marzena Łuczkiwicz