Managerial decision, fuzzy rule-based model, Car Fleet Management, MIS, application requirements, customer segmentation

Anna WALASZEK-BABISZEWSKA*, Michał CHUDZICKI**

FUZZY MODEL FOR THE INFORMATION AND DECISION MAKING SUPPORT SYSTEM FOR THE CFM BRANCH COMPANY

Abstract.

The aim of this work is to present the ways of improvement the management process in the CFM company, by developing and implementing the Management Information System (MIS). In the first chapter the characteristic of Car Fleet Management company is presented. The second chapter describes how MIS could support management of that type of company, as well as general and functional requirements of the computer application. Third chapter presents an example of methodology of solving selected decision-making problem using a fuzzy rule-based model.

1. INTRODUCTION

The Car Fleet Management branch is, without doubt, a novelty on the Polish services market. Owing to favourable legal solutions and growing customer consciousness the branch has been developing very fast and Poland's accession to the European Union further intensifies this tendency as, according to the authors, it propagates in Poland the outsourcing-based car fleet management model, widely utilised throughout Europe. This is confirmed by empirical data (Olivier, 2003), which demonstrate the ratio of the number of cars leased by companies to the number of cars owned by companies in Europe and Poland, which in Europe averages at 40 to 100 whereas in Poland at 5 to 100. There is a certain point in a development of a CFM branch service company when possessing appropriate information and decision-making support system, capable of supporting operating activities as well as providing the management with information of fundamental nature for making strategic decisions, becomes a must. Strategic management which is a multistage process of analysis, planning and management (understood as the stage of strategy implementation) (Stoner et al., 1997), should enable achievement of company's long-term goals by means of information feedback, in a manner adequate to changing environment conditions. In the area of long-term planning, such information system should support a manager in identifying premises that can form basis for making effective strategic decisions with regard to allocation of resources and policy towards

* dr hab. inż. Anna Walaszek-Babiszewska, prof. P. Śl., E-mail: Anna.Walaszek-Babiszewska@polsl.pl
** mgr inż. Michał Chudzicki, E-mail: Michal.Chudzicki@polsl.pl
Silesian University of Technology, Department of Computer Science and Econometrics, Roosevelta 26, 41-800 Zabrze, Poland

customers and contractors. Hence, from the viewpoint of a manager such tool should enable, among others, analysis of the structure of:

- customers for profitability,
- cars for profitability and failure frequency,
- insurance companies for the quality of service,
- cooperating car repair companies for the price levels of spare parts and services.

Considering the need of including the manager's knowledge, the possibility of utilising system's information resources and uncertainty of the information coming from the company's surroundings, the decision-making system includes both the descriptive model and the normative (prescriptive) one. We are able to build such system using the fuzzy sets theory (Bellmann and Zadeh, 1970), (Zadeh, 1975).

In this paper, the authors intend to analyse the customers' structure by means of a dedicated information and decision-making support system in a chosen CFM branch company. Basing on the obtained results the authors are going to present conclusions helpful in formulating the customer service strategy for the analysed company.

2. IMPLEMENTATION OF A SAMPLE APPLICATION FUNCTION

2.1. Verbal description of the problem

Together with the management of the company it has been decided that such analysis of customers' structure should enable breakdown of the customers depending on the cooperation profitability criterion, while the evaluation of cooperation profitability should be based on the following:

- Total revenue in the considered time period – this index shows the actual revenue generated while serving the given customer in the considered time period. The revenue is calculated as the total amount invoiced to the customer.

- Total cost of service in the considered time period. The index shall be calculated basing on the customer's participation, expressed in percentage terms, over the considered operating period of one or more cars and their total operating cost which includes: lease instalments, depreciation or interest on bank loans related to obtaining a car, costs of servicing and repairs, traffic accidents, costs of changing and storing tyres. The customer's participation, expressed in percentage terms, over the car operating period is a ratio of the number of days the customer used the car to the total number of days of the cars availability in the given period of time. Such percentage is also called utilisation.

2.2. Statistical investigations of numerical data

In order to statistically analyse the structure of customers' population, a set of variables is assumed that are essential in the process of decision-making concerning formulating customer service strategy. It is assumed a certain multidimensional random variable $(X_1, X_2,...,X_r) \in \mathbf{X}=R^r$ and a certain finite number J_i of disjoint variation intervals

$$x_{i,j_i} = [x_{i,j_i \min}, x_{i,j_i \max}) \; j = 1,2,..., J_i ; \; i = 1,2,...,r \tag{1}$$

of each variable X_i, such that the probability of a simultaneous event

$$P\left(X_1 \in [x_{1,j_1 \min}, x_{1,j_1 \max}),\ldots, X_r \in [x_{r,j_r \min}, x_{r,j_r \max})\right) = \\ = P_{X_1, X_2,\ldots X_r}(x_{1,j_1}, x_{2,j_2},\ldots, x_{r,j_r}) \tag{2}$$

is constant and is equal to the quotient

$$P_{X_1, X_2,\ldots X_r}(x_{1,j_1}, x_{2,j_2},\ldots, x_{r,j_r}) = \frac{n_{j_1, j_2,\ldots, j_r}}{N} \tag{3}$$

where:

$n_{j_1, j_2,\ldots, j_r}$ - is the number of observed customers whose generate values of variables $X_1,\ldots,X_r$ from the proper intervals,

N – the total number of observations.

Defined in this case rD empirical distribution should fulfil the dependence

$$\sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2}\ldots\sum_{j_r=1}^{J_r}\frac{n_{j_1, j_2,\ldots, j_r}}{N} = 1 \tag{4}$$

There are also marginal (and 2D marginal) empirical probability distributions in the distribution [3].

$$P_{X_1, X_2}(x_{1,j_1}, x_{2,j_2}) = \sum_{X_3,\ldots,X_r}\frac{n_{j_1, j_2,\ldots, j_r}}{N} \tag{5}$$

Variability of revenues generated by customers has been presented in Table 1 and on Figure 1 as well as costs (Table 2, Figure 2). Table 3 presents 2D distribution, expressed by numbers of customers in common intervals of revenues and costs. Table 4 presents opposition of revenues and costs in common intervals.

2.3. Exemplary calculations

The first stage of the customer grouping process was to define the grouping intervals. Due to relatively wide ranges of volume of both revenue and costs generated by customers, it has been jointly decided with the company's management to use a five-interval scale. Experts appointed by the company have arbitrarily defined the limits of the intervals. Table 1 and Figure 1 presents revenues by numbers and amounts in particular intervals. Table 2 and Figure 2 presents costs by numbers and amounts in particular cost intervals. Due to technical conditions both table and figures show the upper limits of intervals in the first column.

Tab.1. Revenues by numbers and amounts in revenues intervals

Revenues intervals, PLN	Numbers of customers	Numbers of customers, %	Revenues, PLN	Revenues, %	Average revenues in intervals, PLN
1 000,00	402	52,62%	188 115,14	0,90%	467,95
25 000,00	295	38,61%	1 466 956,83	6,99%	4 972,74
100 000,00	40	5,24%	2 035 507,53	9,70%	50 887,69
250 000,00	11	1,44%	1 787 069,37	8,52%	162 460,85
3 000 000,00	16	2,09%	15 505 597,76	73,90%	969 099,86
	764	100%	20 983 246,63	100%	

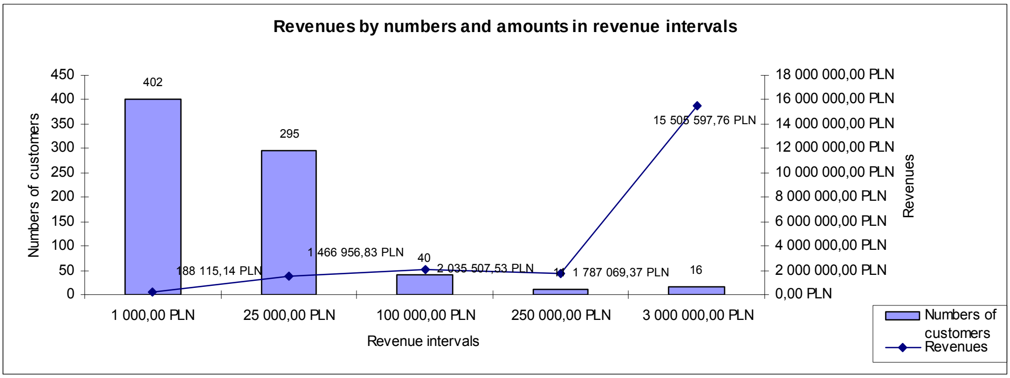

Fig.1. Revenues by numbers and amounts in revenues intervals

Tab.2. Costs by numbers and amounts in cost intervals

Costs intervals, PLN	Numbers of customers	Numbers of customers, %	Costs, PLN	Costs, %	Average costs in intervals, PLN
1 000,00	545	71,34%	140 287,58	1,04%	257,41
25 000,00	165	21,60%	918 297,81	6,79%	5 565,44
100 000,00	34	4,45%	1 587 817,88	11,74%	46 700,53
250 000,00	6	0,79%	999 487,35	7,39%	166 581,22
1 500 000,00	14	1,83%	9 875 404,73	73,04%	705 386,05
	764	100%	13 496 295,35	100%	

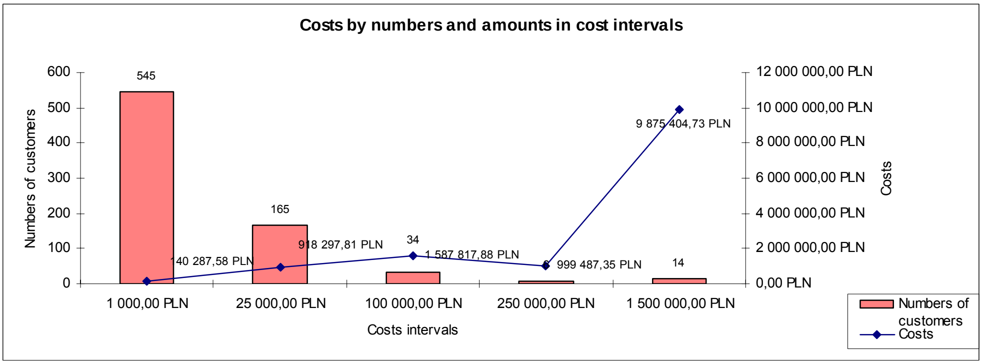

Fig.2. Costs by numbers and amounts in cost intervals

For the sake of global evaluation of profitability of cooperation with customers it is important to study the customers' structure in common intervals. Table 3 presents juxtaposition of numbers of customers in common intervals. Table 4 presents opposition of revenues and costs in common intervals

Tab.3. Juxtaposition of numbers of customers in common intervals of revenues and costs

	Costs intervals, PLN				
Revenues intervals, PLN	1 000,00	25 000,00	100 000,00	250 000,00	1 500 000,00
1 000,00	399	3	0	0	0
25 000,00	146	146	3	0	0
100 000,00	0	15	24	1	0
250 000,00	0	1	5	4	1
3 000 000,00	0	0	2	1	13

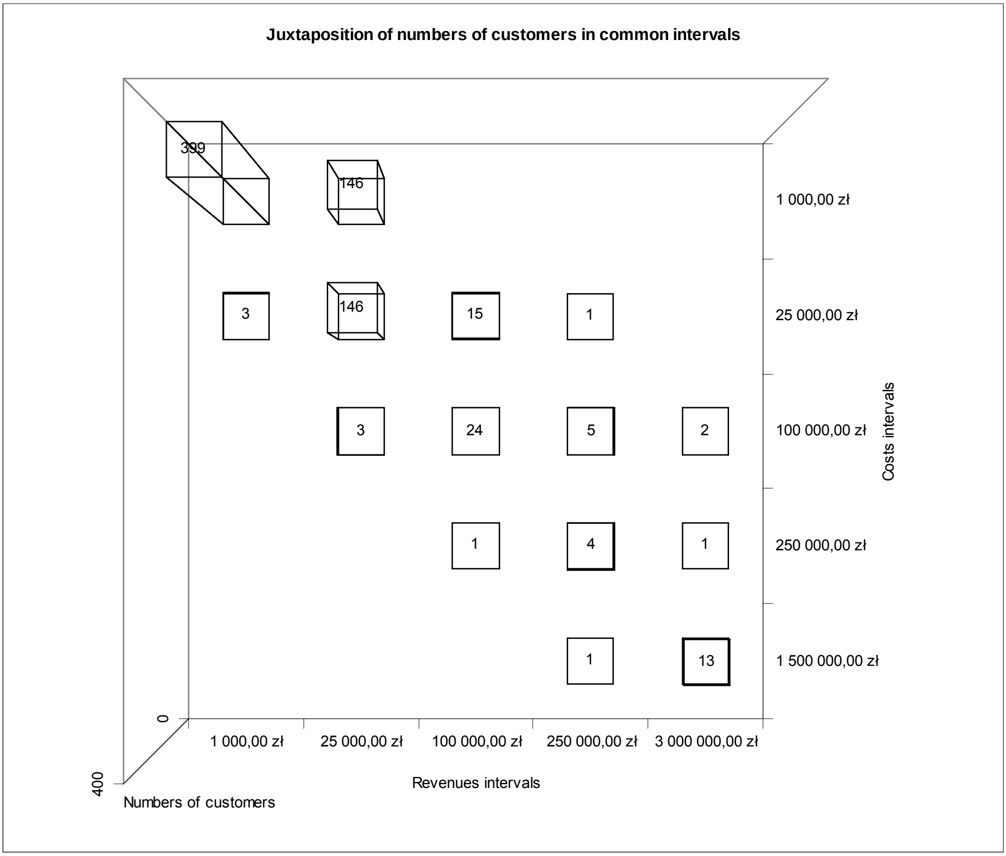

Fig.3. Juxtaposition of numbers of customers in common intervals of revenues and costs

Tab.4. Opposition of revenues and costs in common intervals of revenues and costs, PLN

	Costs intervals, PLN				
Revenues intervals, PLN	1 000,00	25 000,00	100 000,00	250 000,00	1 500 000,00
1 000,00	117 783	-2 102			
25 000,00	222 260	414 059	-40 271		
100 000,00		349 934	258 906	-89 449	
250 000,00		214 951	331 363	192 151	-841 015
3 000 000,00				60 213	4 705 137

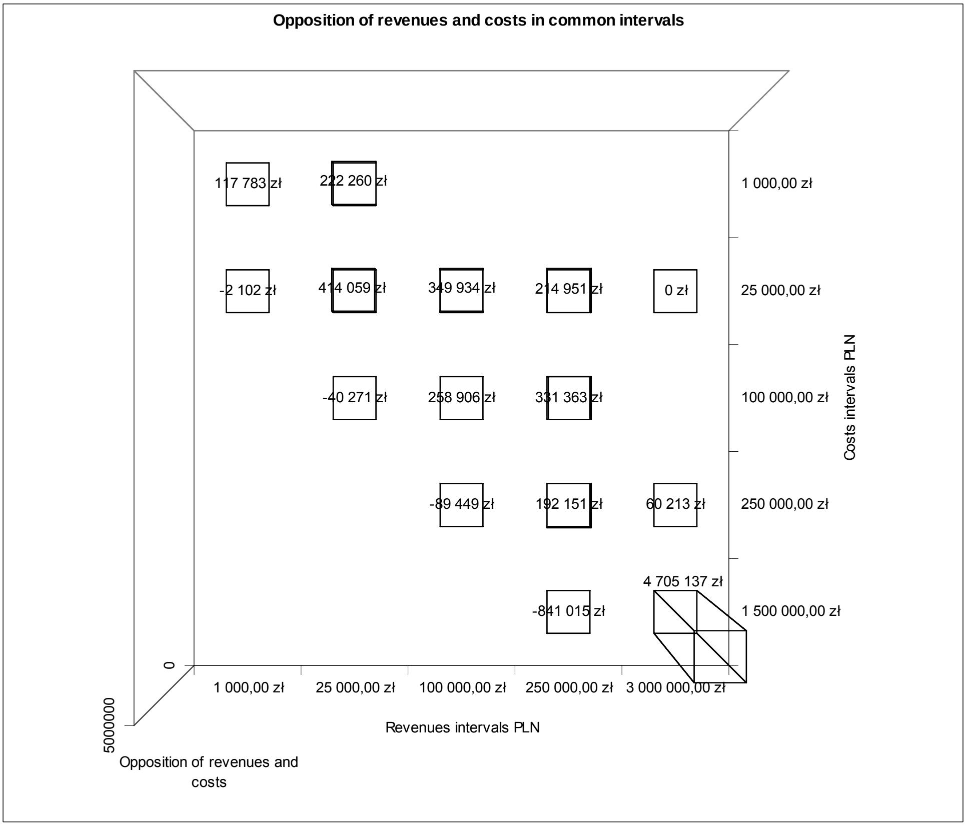

Fig.4. Opposition of revenues and costs in common intervals of revenues and costs, PLN

2.4. Discussion

The following conclusions on profitability of cooperating with customers can be drawn from juxtapositions presented in the previous paragraph:

- Relatively high revenue means that the given customer has frequently used many or expensive cars. This kind of customer is desirable, as the profit he generates is relatively high. Low revenue lets us presume that the customer has used cheap cars rather seldom. This is the most typical kind of customer who only sporadically uses the services of the company.

- Relatively high cost means that the given customer has frequently used many cars or expensive ones or has used them in a way that has generated high costs. Low costs mean the customer has seldom used cheap cars. Unequivocal customer evaluation is not possible by looking only at costs.

- Comparison of revenue and servicing cost. In the event of high revenue and high servicing cost we can suppose we are dealing with a big and important customer that uses a number of cars. Such customer is a good customer that guarantees steady revenue at a high level. However, disproportionately high cost tells us that the customer generates high operating costs (traffic accidents, failures). High revenue combined with low costs suggests that customer uses expensive cars and utilises them in a way that does not cause additional costs. This kind of customer is most desirable for the company. Low revenue combined with low costs indicates a typical small customer, sporadically using the company's services. Low revenue and high cost mean we are dealing with a customer that ineffectively uses a small number of cars. Such customers should be avoided.

3. FUZZY MODEL FOR THE DECISION-MAKING TASK

3.1. General form of the decision-making model

When constructing an information processing system such as a classifier of the costumer population, two kinds of information are available. One is numerical data from observations and its statistical analysis and the other is linguistic information from human experts. In this part of the paper a fuzzy rule-based model is proposed for constructing a classification system. The model can be seen as Mamdani fuzzy model (Hellendorn et al., 1997). The model structure, input and output variables, the partition of the space of variables are related to the expert knowledge. Statistical analysis of numerical data is helpful for the validation of rules.

In general, the model takes into account the revenue-cost combination of the process parameters to distinguish some groups of customers, interesting from manager's point of view, e.g.:

w_1: if *X_1 is high* and *X_2 is low* then *C belongs to class 1* (the best)
w_2: if *X_1 is high* and *X_2 is middle* then *C belongs to class 2* (good)
w_3: if *X_1 is high* and *X_2 is high* then *C belongs to class 3* (high risk) (6)
…………………………………………………………………
w_9: if *X_1 is low* and *X_2 is low* then *C belongs to class 9* (typical)

where

X_1, X_2 are considered variable, revenue and cost, respectively;
'low', 'middle', 'high' are the linguistic values of the variables X_1, X_2 (Zadeh, 1975) whose are represented by fuzzy sets and membership coefficients;
$w_1, \ldots, w_9$ – weights of rules, equal to a probability of fuzzy events
(*X_1 is high*)*and* (*X_2 is low*),
………………………….
(*X_1 is low*)*and*(*X_2 is low*),
respectively.

Probability *P(X_1 is high)* of the fuzzy event, representing the linguistic value of variable X_1 , according to Zadeh's definition (Zadeh, 1968), is equal to:

$$P(X_1\, is\, high) = \sum_i \mu_{high}(a_i)\, p_{X_1}(a_i) \tag{7}$$

and a joint probability of the fuzzy event (*X_1 is high*) $\cap$ (*X_2 is low*) can be calculated as follows:

$$P((X_1\, is\, high) \cap (X_2\, is\, low)) = \sum_{i,j} \mu_{high}(a_i)\, \mu_{low}(b_j)\, p_{X_1,X_2}(a_i, b_j) \tag{8}$$

where the probabilities $p_{X_1,X_2}(a_i,b_j)$, $p_{X_1}(a_i)$ are real numbers, calculated according to (1) – (5) and $\mu_{high}(a_i)$, $\mu_{low}(b_j)$ are membership functions of the respective fuzzy sets.

3.2. Linguistic variables of the process

It is assumed that linguistic variable 'revenues' is representing in the space X, by the set of linguistic values L(X):= {low, middle, high}.

Fuzzy sets A_1, A_2, A_3 representing particular linguistic values, are described in the space X as follows:

$$A_1 := \text{'low'} \quad \mu_{A_1}(x) = \begin{Bmatrix} 1 \text{ for x } \in \Delta x_1 \\ 0,5 \text{ for x } \in \Delta x_2 \end{Bmatrix}$$

$$A_2 := \text{'middle'} \quad \mu_{A_2}(x) = \begin{Bmatrix} 0,5 \text{ for x} \in \Delta x_2 \\ 1 \text{ dla for} \in \Delta x_3 \\ 0,5 \text{ for x} \in \Delta x_4 \end{Bmatrix} \tag{9}$$

$$A_3 := \text{'high'} \quad \mu_{A_3}(x) = \begin{Bmatrix} 0,5 \text{ for x} \in \Delta x_4 \\ 1 \text{ for x} \in \Delta x_5 \end{Bmatrix}$$

where Δx_i i=1,…,5 are intervals shown in the Table 1.

Defined in this case membership functions fulfil the dependence:

$$\forall \text{x} \in \Delta x_i \quad \sum\nolimits_{\text{m}=1,2,3} \mu_{A_m}(\text{x} \in \Delta x_i) = 1 \quad \text{i} = 1,\, 2,\, ...,5. \tag{10}$$

Also it is assumed, that linguistic variable 'costs' is representing in the space Y, by the set of linguistic values L(Y) : ={low, middle, high}.

Fuzzy sets B_1, B_2, B_3 representing particular linguistic value, was described in the space Y as follows:

$$B_1 := \text{'low'} \quad \mu_{B_1}(y) = \begin{Bmatrix} 1 \text{ for y} \in \Delta y_1 \\ 0,5 \text{ for y} \in \Delta y_2 \end{Bmatrix}$$

$$B_2 := \text{'middle'} \quad \mu_{B_2}(y) = \begin{Bmatrix} 0,5 \text{ for y} \in \Delta y_2 \\ 1 \text{ for y} \in \Delta y_3 \\ 0,5 \text{ for y} \in \Delta y_4 \end{Bmatrix} \tag{11}$$

$$B_3 := \text{'high'} \quad \mu_{B_3}(y) = \begin{Bmatrix} 0,5 \text{ for y} \in \Delta y_4 \\ 1 \text{ for y} \in \Delta y_5 \end{Bmatrix}$$

Defined in this case membership functions fulfil the dependence:

$$\forall y \in \Delta y_j \quad \sum\nolimits_{\text{n}=1,2,3} \mu_{B_n}(\text{y} \in \Delta y_j) = 1, \quad \text{j} = 1,2,...,5. \tag{12}$$

Linguistic vector variable z, named 'customers classification', defined in Z=X × Y takes the set of linguistic values L(Z):={'small customers', 'low risk customers', 'preferred middle customers', 'middle customers', 'middle risk customers', 'top customers', 'very good customers', 'good customers'}. The fuzzy relations represent particular linguistic values, as follows:

$$
\begin{aligned}
&A_1 \cap B_1 := \text{'small customers'}\\
&A_1 \cap B_2 := \text{'low risk customers'}\\
&A_1 \cap B_3 := \text{empty set,}\\
&A_2 \cap B_1 := \text{'preferred middle customers'}\\
&A_2 \cap B_2 := \text{'middle customers'}\\
&A_2 \cap B_3 := \text{'middle risk customers'}\\
&A_3 \cap B_1 := \text{'top customers'}\\
&A_3 \cap B_2 := \text{'very good customers'}\\
&A_3 \cap B_3 := \text{'good customers'}
\end{aligned}
\tag{13}
$$

where the relation $A_l \cap B_k$ is determined as follows:

$$\mu_{A_l \cap B_k}(a_i, b_j) = \mu_{A_l}(a_i)\mu_{B_k}(b_j) \tag{14}$$

3.3. Probability of fuzzy events

Empirical probability distribution of two linguistic variables of a simultaneous event will be calculated according to the formula (Zadeh, 1968), (Walaszek-Babiszewska, 2003):

$$P(A_l \cap B_k) = \sum_{i,j} p_{ij}(x \in \Delta x_i, y \in \Delta y_j)\, \mu_{A_l}(\Delta x_i)\, \mu_{B_k}(\Delta y_j) \tag{15}$$

Where the probability of a simultaneous event $p_{ij}(x \in \Delta x_i, y \in \Delta y_j)$, according to (3), is constant and is equal to the quotient:

$$p_{ij}(x \in \Delta x_i, y \in \Delta y_j) = \frac{n_{ij}}{n} \tag{16}$$

n_{ij} – is the number of observed customers whose generate values of variables $x_1, \ldots, x_5$ and $y_1, \ldots, y_5$ from the proper intervals,

n – the total number of customers.

Using data from the Table 3 and the formula (15) it is possible to calculate the probability, that selected customer belongs to the common space, mapped by fuzzy sets $A_i \cap B_j$. For example:

$$
\begin{aligned}
&p(A_1 \cap B_1) = p_{11}(x \in \Delta x_1, y \in \Delta y_1) \times \mu_{A_1}(\Delta x_1) \times \mu_{B_1}(\Delta y_1) + \\
&p_{12}(x \in \Delta x_1, y \in \Delta y_2) \times \mu_{A_1}(\Delta x_1) \times \mu_{B_1}(\Delta y_2) + p_{21}(x \in \Delta x_2, y \in \Delta y_1) \times \mu_{A_1}(\Delta x_2) \times \mu_{B_1}(\Delta y_1) + \\
&p_{22}(x \in \Delta x_2, y \in \Delta y_2) \times \mu_{A_1}(\Delta x_2) \times \mu_{B_1}(\Delta y_2) = \\
&= 0{,}5223 \times 1 \times 1 + 0{,}19 \times 0{,}5 \times 1 + 0{,}004 \times 1 \times 0{,}5 + 0{,}19 \times 0{,}5 \times 0{,}5 = 0{,}6675
\end{aligned}
\tag{17}
$$

where × is the symbol of an algebraic multiplication.

Similarly calculated probabilities for all relationships (13) have been presented in table 5.

Defined in this case 2D empirical distribution of linguistic variables fulfils the dependence:

$$\sum_{i,j=1,2,3} P\,(A_i \cap B_j) = 1 \tag{18}$$

Tab.5. Empirical probability distribution of the simultaneous event that selected customer belongs to common space of fuzzy sets $A_i \cap B_j$.

	B_1	B_2	B_3
A_1	P('small customers')= 0,6675	P('low risk customers')= 0,0517	P(empty set) = 0
A_2	P('preferred middle customers') = 0,1535	P('middle customers') = 0,0965	P('middle risk customers') = 0,0026
A_3	P('top customers') = 0,0003	P('very good customers')= 0,0082	P('good customers') = 0,0196

3.4. Rule-based classification fuzzy model

Rule based model of the customer classification expressed in fuzzy categories has the form:

R^i w_i(IF x is A_k AND y is B_l THEN z is C_i)

where A_k ,B_l are fuzzy sets of the revenue-cost combination, C_i is a name of i-th class of customers, distinguished from the manager's point of view, w_i i=1,…,9 are the weights of rules calculated in the above paragraph.

The same model can be expressed in linguistic categories as follows:

R^1: 0,6675(IF x is 'low' AND y is 'low' THEN z is 'small customers')
R^2: 0,0517(IF x is 'low' AND y is 'middle' THEN z is 'low risk customers')
R^4: 0,1535(IF x is 'middle' AND y is 'low' THEN z is 'preferred middle customers')
R^5: 0,0965(IF x is 'middle' AND y is 'middle' THEN z is 'middle customers') (19)
R^6: 0,0026(IF x is 'middle' AND y is 'high' THEN z is 'middle risk customers')
R^7: 0,0003(IF x is 'high' AND y is 'low' THEN z is 'top customers')
R^8: 0,0082(IF x is 'high' AND y is 'middle' THEN z is 'very good customers')
R^9: 0,0196(IF x is 'high' AND y is 'high' THEN z is 'good customers')

4. CONCLUSIONS AND FUTURE WORK

Developed decision support system includes also another modules for the classification of others data, i.e.: cars by marks, models and segments, insurance companies etc.

We can check also a stability of the classification, taking into account and comparing the probability of particular consumer segments, calculated on the base of the data sets from the next year $t=t_{k+1}$, e.g.:

$P(A_1 \cap B_1 :=$ 'small customers') >0,6675
$P(A_1 \cap B_2 :=$ 'low risk customers') <0,0517
$P(A_2 \cap B_1 :=$ 'preferred middle customers') >0,1535
$P(A_2 \cap B_2 :=$ 'middle customers') <0,0965 (20)
$P(A_2 \cap B_3 :=$ 'middle risk customers') <0,0026
$P(A_3 \cap B_1 :=$ 'top customers') >0,0003
$P(A_3 \cap B_2 :=$ 'very good customers') >0,0082
$P(A_3 \cap B_3 :=$ 'good customers') >0,0196

References

[1] BELLMANN R.E., ZADEH L.A.: *Decision making in a fuzzy environment*. Management Science, 1970, nr 17, 141-164.

[2] KACPRZYK J.: *Fuzzy sets in system analysis (in Polish)*, PWN, Warsaw 1986.

[3] HELLENDOORN H., DRIANKOV D.: *Fuzzy model identification*. Selected Approaches, Springer-Verlag, Berlin, Heidelberg 1997.

[4] OLIVIER B.: *Car politics in a company* (in Polish). Flota. Magazyn kadry zarządzającej flotą samochodowa, 2003, nr 5, 26-29.

[5] STONER J.A.F., WANKEL Ch.: *Management* (in Polish), PWE, Warsaw 1997.

[6] WALASZEK-BABISZEWSKA A.: *Fuzzy probability for modelling of particle preparation processes, in: Intelligence in a Small World*; (Selected Papers from IPMM-2003 The Fourth Int. Conf. Intelligent Processing and Manufacturing of Materials, Nanotechnology for the 21st Century. Sendai, Japan, May 2003); Meech J. A., Kawazoe Y., Kumar V., Maguire J. F. (Eds.); pp. 119-129; DEStech Publ. Inc. Lancaster,USA 2005 (+CD-ROM).

[7] ZADEH L.A.: *Probability measures of fuzzy events*. J. Math. Anal. Appl., 1968, nr 23, 421-427.

[8] ZADEH L.A.: *The concept of a linguistic variable and its application to approximate reasoning*. Information Sciences, Part I: 8 (1975) 199-240.