Selected Methods of Combining Classifiers, when Predictions Are Stored in Probability Vectors, in a Dispersed Decision-Making System

Malgorzata Przybyla-Kasperek

University of Silesia, Institute of Computer Science,
Bedzinska 39, 41-200 Sosnowiec, Poland
malgorzata.przybyla-kasperek@us.edu.pl
http://www.us.edu.pl

Abstract. Issues that are related to decision making that is based on dispersed knowledge are discussed in the paper. A system, that was proposed in the article [12], is used in this paper. In the system the process of combining classifiers in coalitions is very important and negotiation is applied in the clustering process. The main aim of the article is to compare the results obtained using five different methods of conflict analysis in the system. All these methods are used if the individual classifiers generate probability vectors over decision classes. The most popular methods are considered: a sum, a product, a median, a maximum and a minimum rules. In the paper, tests, which were performed on data from the UCI repository, are presented. The best methods in a particular situation are indicated. It was found out that some methods do not generate satisfactory results when there are dummy agents in a dispersed data set. That is there are undecided agents who assign the same probability value to many different decision values.

Key words: decision support system, dispersed knowledge, conflict analysis, sum rule, product rule, median rule, maximum rule, minimum rule

1 Introduction

The effectiveness of decision making based on dispersed knowledge is an increasingly vital issue. When important economic, medical, enterprise management, business and risk assessment or political issues are to be resolved, groups rather than individuals are employed to make high-quality decisions. Compared to individual decision makers, groups have access to more and a broader range of information, which is distributed among group members. So making decisions based on dispersed knowledge is more difficult, demanding and requires new methods of inference. In this paper an approach proposed in the article [12] is used to make decisions on the basis of dispersed knowledge. The aim of the paper is to investigate the use of five selected conflict analysis methods in the system with dynamically generated clusters. The problem of conflict analysis arises because the inference is being conducted in groups of knowledge bases. Five methods known from the literature [3, 5, 6] were used to conflict analysis: the sum rule, the product rule, the median rule, the maximum rule and the minimum rule.

Several attempts have already been made to solve the problem of decision making that is based on dispersed knowledge, group decisions and negotiations. In the paper [1], a group decision-making (GDM) problem in which linguistic information is used is considered. The method, which is based on granular computing and pairwise comparisons, was used to consider GDM situations that were defined in heterogeneous contexts, that is, situations in which the experts have different backgrounds and levels of knowledge about the problem. A discussion of pairwise comparisons of objects was presented in the paper [4]. The concept of distributed decision-making is widely discussed in the papers [13, 14]. In addition, the problem of using distributed knowledge is discussed in many other papers [2, 15, 16]. In the paper [18], an approach was proposed in which many classifying agents are generated by using fast approximation heuristics, after which a classification system is constructed by selecting the optimal subset of agents. This paper describes a different approach to the global decision-making process. We assume that the set of local knowledge bases that contain information from one domain is pre-specified. The only condition which must be satisfied by the local knowledge bases is to have common decision attributes.

2 A Brief Overview of a Dispersed Decision-Making System

The concept of a dispersed decision-making system is being considered by the author for several years. In the first stage of studies the considerations were directed to a system with a static structure [10, 17]. In recent papers a system with a dynamic structure has been proposed [11, 12]. During the construction of this system's structure a negotiation stage is used. The main assumptions, notations and definitions of the system are described below. A detailed discussion can be found in the paper [12].

We assume that the knowledge is available in a dispersed form, which means in a form of several decision tables. Each local knowledge base is managed by one agent, which is called a resource agent. We call ag in $Ag = \{ag_1, \ldots, ag_n\}$ a resource agent if he has access to resources represented by a decision table $D_{ag} := (U_{ag}, A_{ag}, d_{ag})$, where U_{ag} is a finite nonempty set called the universe; A_{ag} is a finite nonempty set of conditional attributes, V^a_{ag} is a set of attribute a values; d_{ag} is referred to as a decision attribute, V^d_{ag} is called the value set of d_{ag}. We want to designate homogeneous groups of resource agents. The agents who agree on the classification for a test object into the decision classes will be combined in the group. It is realized in two steps. At first initial coalitions are created. Then the negotiation stage is implemented. These two steps are based on the test object classification carried out by the resource agents. For each agent the classification is represented as a vector of values, whose dimension is equal to the number of decision classes. This vector will be defined on the basis of certain relevant objects. That is the objects from the decision tables of agents that carry the greatest similarity to the test object. From decision table of resource agent $D_{ag}, ag \in Ag$ and from each decision class, the smallest set containing at least m_1 objects for which the values of conditional attributes bear the greatest similarity to the test object is chosen. The value of the parameter m_1 is selected experimentally. Then for each resource agent $i \in \{1, \ldots, n\}$ and the test object x, a c-dimensional vector $[\bar{\mu}_{i,1}(x), \ldots, \bar{\mu}_{i,c}(x)]$ is generated. The value $\bar{\mu}_{i,j}(x)$ is equal to the average value of the similarity of the test object to the relevant objects of agent ag_i, belonging to the decision class v_j. In the experimental part of this paper the Gower similarity measure [11] was used. This measure enables the analysis of data sets that have qualitative, quantitative and binary attributes. On the basis of the vector of values defined above, a vector of the rank is specified. The vector of rank is defined as follows: rank 1 is assigned to the values of the decision attribute that are taken with the maximum level of certainty. Rank 2 is assigned to the next most certain decisions, etc. Proceeding in this way for each resource agent $ag_i, i \in \{1, \ldots, n\}$, the vector of rank $[\bar{r}_{i,1}(x), \ldots, \bar{r}_{i,c}(x)]$ will be defined. In order to create clusters of agents, relations between the agents are defined. These definitions were taken from the papers of Pawlak [8, 9]. Relations between agents are defined by their views on the classification of the test object x to the decision class. We define the function $\phi^x_{v_j}$ for the test object x and each value of the decision attribute $v_j \in V^d$; $\phi^x_{v_j} : Ag \times Ag \to \{0, 1\}$ $\phi^x_{v_j}(ag_i, ag_k) = \begin{cases} 0 & \text{if } \bar{r}_{i,j}(x) = \bar{r}_{k,j}(x) \\ 1 & \text{if } \bar{r}_{i,j}(x) \neq \bar{r}_{k,j}(x) \end{cases}$ where $ag_i, ag_k \in Ag$. We also define the intensity of conflict between agents using a function of the distance between agents. We define the distance between agents ρ^x for the test object x: $\rho^x : Ag \times Ag \to [0, 1]$, $\rho^x(ag_i, ag_k) = \frac{\sum_{v_j \in V^d} \phi^x_{v_j}(ag_i, ag_k)}{card\{V^d\}}$, where $ag_i, ag_k \in Ag$.

Definition 1. *Let p be a real number, which belongs to the interval $[0, 0.5)$. We say that agents $ag_i, ag_k \in Ag$ are in a friendship relation due to the object x, which is written $R^+(ag_i, ag_k)$, if and only if $\rho^x(ag_i, ag_k) < 0.5 - p$. Agents $ag_i, ag_k \in Ag$ are in a conflict relation due to the object x, which is written $R^-(ag_i, ag_k)$, if and only if $\rho^x(ag_i, ag_k) > 0.5 + p$. Agents $ag_i, ag_k \in Ag$ are in a neutrality relation due to the object x, which is written $R^0(ag_i, ag_k)$, if and only if $0.5 - p \leq \rho^x(ag_i, ag_k) \leq 0.5 + p$.*

By using the relations defined above we can create groups of resource agents, which are not in conflict relation. The initial cluster due to the classification of object x is the maximum, due to the inclusion relation, subset of resource agents $X \subseteq Ag$ such that $\forall_{ag_i, ag_k \in X} \; R^+(ag_i, ag_k)$. In the second stage of clustering, limitations imposed on compatibility of agents are relaxed. We assume that during the negotiation, agents put the greatest emphasis on compatibility of ranks assigned to the decisions with the highest ranks. We define the function ϕ^x_G for the test object x; $\phi^x_G : Ag \times Ag \to [0, \infty)$ $\phi^x_G(ag_i, ag_j) = \frac{\sum_{v_l \in Sign_{i,j}} |\bar{r}_{i,l}(x) - \bar{r}_{j,l}(x)|}{card\{Sign_{i,j}\}}$ where $ag_i, ag_j \in Ag$ and $Sign_{i,j} \subseteq V^d$ is the set of significant decision values for the pair of agents ag_i, ag_j. In the set $Sign_{i,j}$ there are the values of the decision, which the agent ag_i or agent ag_j gave the highest rank. During the negotiation stage, the intensity of the conflict between the two groups of agents is determined by using the generalized distance. The generalized distance between agents for the test object x is denoted by ρ^x_G; $\rho^x_G : 2^{Ag} \times 2^{Ag} \to [0, \infty)$. The value of the generalized distance function for two sets of agents X and Y is equal to the average value of the function ϕ^x_G for each pair of agents ag, ag' belonging to the set $X \cup Y$. This value can be interpreted as the average difference of the ranks assigned to significant decisions within the combined group of agents consisting of the sets X and Y. For each agent ag that has not been included to any initial clusters, the generalized distance value is determined for this agent and all initial clusters, with which the agent ag is not in a conflict relation and for this agent and other agents without coalition, with which the agent ag is not in a conflict relation. Then the agent ag is included to all initial clusters, for which the generalized distance does not exceed a certain threshold, which is set by the system's user. Also agents without coalition, for which the value of the generalized distance function does not exceed the threshold, are combined into a new cluster. After completion of the second stage of the process of clustering we get the final form of clusters. For each cluster, a superordinate agent is defined, which is called a synthesis agent, as_j, where j- number of cluster. As_x is a finite set of synthesis agents defined for the clusters that are dynamically generated for test object x. Next, an approximated method of the aggregation of decision tables have been used to generate decision tables for synthesis agents (see [10–12] for more details). Based on these aggregated decision tables global decisions are taken using the methods of conflict analysis.

3 Methods of Conflict Analysis

In this article, we use five different methods of conflict analysis: the sum rule, the product rule, the median rule, the maximum rule and the minimum rule. These are well known and commonly used methods in group decision-making problems. They are discussed by various authors [3, 5–7]. The methods, that are used in this article, are simple, have low computational complexity and are easy to implement. These methods require no training and all classifiers are treated equally. In some applications, it may be undesirable, because the methods do not take into account the differences in the individual classifier capabilities. But, as we know from the literature, it is somewhat surprising to see how well these simple aggregation rules compete with the more sophisticated ones. The novelty that is proposed in this article involves the use of these five methods in a dispersed decision-making system that was briefly described above. All these methods are used if the individual classifiers generate vectors of probabilities instead of unique class choices. Therefore at first, on the basis of each aggregated decision table a vector of probabilities is generated. A c-dimensional vector of values $[\mu_{j,1}(x), \ldots, \mu_{j,c}(x)]$ is generated for each j-th cluster, where c is the number of all of the decision classes. This vector will be defined on the basis of relevant objects. From each aggregated decision table and from each decision class, the smallest set containing at least m_2 objects for which the values of conditional attributes bear the greatest similarity to the test object is chosen. The value of the parameter m_2 is selected experimentally. The value $\mu_{j,i}(x)$ is equal to the average value of the similarity of the test object to the relevant objects form j-th aggregated decision table, belonging to the decision class v_i. In this way, for each cluster the vector of probabilities is generated.

In the paper [6], it was proposed that the classifier outputs can be organized in a decision profile (DP) as the matrix. This is very clear and transparent way of classifiers outputs presentation. The decision profile is a matrix with dimensions $card\{As_x\} \times c$, where As_x is a finite set of synthesis agents defined for the test object x and c is the number of all of the decision classes. The decision profile is defined as follows

$$DP(x) = \begin{bmatrix} \mu_{1,1}(x) & \cdots & \mu_{1,i}(x) & \cdots & \mu_{1,c}(x) \\ \cdots & & & & \\ \mu_{card\{As_x\},1}(x) & \cdots & \mu_{card\{As_x\},i}(x) & \cdots & \mu_{card\{As_x\},c}(x) \end{bmatrix}$$

The j-th row of the matrix saves the output of j-th synthesis agents and the i-th column of the matrix saves support from agents As_x for decision class i.

The sum rule

The sum rule consists in the designation for each decision class the sum of the probability values assigned to this class by each cluster. The set of decisions taken by the dispersed system is the set of classes which have the maximum of these sums. Thus, the set of global decisions that are generated using the sum rule is defined as follows $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \arg\max_{i\in\{1,\ldots,c\}} \left\{ \sum_{j=1}^{card\{As_x\}} \mu_{j,i}(x) \right\}$.

The product rule

In the product rule for each decision class the product of the probability values is determined. The set of decisions taken by the dispersed system is the set of classes which have the maximum of these products $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \arg\max_{i\in\{1,\ldots,c\}} \left\{ \prod_{j=1}^{card\{As_x\}} \mu_{j,i}(x) \right\}$. The approach, that is used in this paper, has a small modification. The product rule is very sensitive to the most pessimistic prediction result of the base classifier. The worst is the situation in which one of the classifiers generate, for several decision classes, probability equal to 0. This situation is called a veto mechanism - one classifier is decisive. To eliminate this drawback, a rule was adopted, that if the probability for the decision class is equal to 0, then the values of probabilities are multiplied by 10^{-3} instead of 0.

The median rule

In the median rule for each decision class the median value of the probability values is determined. The median can be found by arranging all the values from lowest value to highest value and picking the middle one. If there is an even number of values, the median is defined to be the mean of the two middle values. The set of decisions taken by the dispersed system is the set of classes which have the maximum of these medians $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \underset{i\in\{1,\ldots,c\}}{\arg\max} \left\{ \mathrm{med}_{\mathrm{j}\in\{1,\ldots,\mathrm{card}\{\mathrm{As_x}\}\}}\ \mu_{\mathrm{j,i}}(\mathrm{x}) \right\}$.

The maximum rule and the minimum rule

The maximum rule and the minimum rule consist in the designation for each decision class the maximum or the minimum value of the probability values assigned to this class by each cluster. The set of decisions taken by the dispersed system is the set of classes which have the maximum of these values. Thus, the sets of global decisions that are generated using these methods are defined as follows: the maximum rule $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \underset{i\in\{1,\ldots,c\}}{\arg\max} \left\{ \max_{j\in\{1,\ldots,card\{As_x\}\}} \mu_{j,i}(x) \right\}$, the minimum rule $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \underset{i\in\{1,\ldots,c\}}{\arg\max} \left\{ \min_{j\in\{1,\ldots,card\{As_x\}\}} \mu_{j,i}(x) \right\}$.

Example 1. Consider a dispersed decision-making system in which there are three decision classes $V^d = \{v_1, v_2, v_3\}$ and, for a given test object, the set of synthesis agents consists of five agents. We assume that the decision profile is as follows:

$$DP(x) = \begin{bmatrix} 0.5 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.1 \\ 0.6 & 0.4 & 0.2 \\ 0.4 & 0.5 & 0.4 \\ 0.5 & 0.6 & 0.3 \end{bmatrix}$$

Applying each of the methods columnwise, we obtain as a result the vectors and the sets of global decisions. The sum rule: $[2.2, 2.2, 1.3]$, $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \{v_1, v_2\}$; the product rule: $[0.012, 0.014, 0.001]$, $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \{v_2\}$; the median rule: $[0.5, 0.4, 0.3]$, $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \{v_1\}$; the maximum rule: $[0.6, 0.6, 0.4]$, $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \{v_1, v_2\}$; the minimum rule: $[0.2, 0.3, 0.1]$, $\hat{d}_{WSD^{dyn}_{Ag}}(x) = \{v_2\}$. As can be seen, these methods provide different results, and sometimes more than one decision value is generated. This ambiguity of the methods are discussed in more detail in the next section.

4 Results of Experiments

The aim of the experiments is to compare the quality of the classification made by the dispersed decision-making system using five the most popular methods of combining the prediction's results when vectors of probabilities are generated by the base classifiers. The sum, the product, the median, the maximum and the minimum rules were considered. For the experiments the following data, which are in the UCI repository [19], were used: Soybean Data Set and Vehicle Silhouettes data set. In order to determine the efficiency of inference of the dispersed decision-making system with respect to the analyzed data, the Vehicle Silhouettes data set was divided into two disjoint subsets: a training set and a test set. The Soybean data set is available on the UCI repository website in a divided form: a training and a test set. A numerical summary of the data sets is as follows: Soybean: # The training set - 307; # The test set - 376; # Conditional - 35; # Decision - 19; Vehicle Silhouettes: # The training set - 592; # The test set - 254; # Conditional - 18; # Decision - 4. Because the available data sets are not in the dispersed form, in order to test the dispersed decision-making system the training set was divided into a set of decision tables. Divisions with a different number of decision tables were considered. For each of the data sets used, the decision-making system with five different versions (with 3, 5, 7, 9 and 11 decision tables) were considered. For these systems, we use the following designations: WSD^{dyn}_{Ag1} - 3 decision tables; WSD^{dyn}_{Ag2} - 5 decision tables; WSD^{dyn}_{Ag3} - 7 decision tables; WSD^{dyn}_{Ag4} - 9 decision tables; WSD^{dyn}_{Ag5} - 11 decision tables. Note that the division of the data set was not made in order to improve the quality of the decisions taken by the decision-making system, but in order to store the knowledge in a dispersed form. We consider the situation, that is very common in life, in which data from one domain are collected by different units as separate knowledge bases. For each data set we have 5 versions of the dispersion, therefore it can be said that 10 different dispersed data set were used for experiments. During the experiments, it turned out that not all analyzed methods of combining individual classifiers, generate unequivocal decision. Sometimes when making global decisions ties occur. It was noted that this situation occurs particularly for the Soybean data set, when we are using the maximum and the minimum rule. Because we want to observe and analyze such cases, an additional method of ties resolving were not applied. But the appropriate classification measures were applied, which are adequate to this situation. The measures of determining the quality of the classification are: *estimator of classification error* e in which an object is considered to be properly classified if the decision class used for the object belonged to the set of global decisions generated by the system; *estimator of classification ambiguity error* e_{ONE} in which object is considered to be properly classified if only one, correct value of the decision was generated to this object; *the average size of the global decisions sets* $\overline{d}_{WSD_{Ag}^{dyn}}$ generated for a test set. In the description of the results of experiments for clarity some designations for algorithms and parameters have been adopted: m_1 - parameter which determines the number of relevant objects that are selected from each decision class of the decision table and are then used in the process of cluster generation; p - parameter which occurs in the definition of friendship, conflict and neutrality relations; $A(m)$ - the approximated method of the aggregation of decision tables; $C(m_2)$ - the method of conflict analysis (the sum rule, the product rule, the median rule, the maximum rule or the minimum rule), with parameter which determines the number of relevant objects that are used to generate the probability vectors. The process of parameters optimization was carried out as follows. A series of tests for different parameter values were performed: $m_1 \in \{1, 6, 11, 16, 20\}$, $m, m_2 \in \{1, \ldots, 10\}$ and $p \in \{0.05, 0.15\}$. Thus, for each of the ten considered dispersed systems, 1000 tests were conducted ($1000 = 5 \cdot 10 \cdot 10 \cdot 2$). From all of the obtained results, one was selected that guaranteed a minimum value of estimator of classification error (e), while maintaining the smallest possible value of the average size of the global decisions sets ($\overline{d}_{WSD_{Ag}^{dyn}}$). In tables presented below the best results, obtained for optimal values of the parameters, are given.

The results of the experiments with the Soybean data set are presented in Table 1. In the table the following information is given: the name of dispersed decision-making system (System); the selected, optimal parameter values (Parameters); the algorithm's symbol (Algorithm); the three measures discussed earlier $e, e_{ONE}, \overline{d}_{WSD_{Ag}^{dyn}}$; the time t needed to analyze a test set expressed in minutes. As can be seen, for the Soybean data set, unequivocal decisions are generated by the sum and the product rules. Which means that the average size of the global decisions sets is equal to 1. In the case of the median rule depending on chosen parameters, we get a unequivocal decision or a set of decisions having on average 1.5 decisions. For the maximum and the minimum rules regardless of the selected values of the parameters we get quite large set of decisions, sometimes average size is close to the value 2. For other analyzed values of the parameters even larger average number was observed. This ambiguity causes that these two methods are not very useful in the case of the Soybean data set. On the basis of detailed analysis of vectors of probabilities generated by the individual classifiers, it was concluded that the reason of this situation is that for the Soybean data there is a lot of dummy agents. That is there are undecided agents who assign the same probability value to many different decision values. Figure 1 shows a graphical comparison of the estimator of classification error for different dispersed systems of the Soybean data set.

Table 1. Summary of experiments results with the Soybean data set

The sum rule

System	Parameters	Algorytm	e	e_{ONE}	$\overline{d}_{WSD^{dyn}_{Ag}}$	t
WSD^{dyn}_{Ag1}	$m_1 = 6, p = 0.05$	$A(3)C(3)$	0.088	0.088	1	0.05
WSD^{dyn}_{Ag2}	$m_1 = 11, p = 0.15$	$A(2)C(5)$	0.093	0.093	1	0.07
WSD^{dyn}_{Ag3}	$m_1 = 11, p = 0.15$	$A(3)C(9)$	0.096	0.096	1	0.11
WSD^{dyn}_{Ag4}	$m_1 = 1, p = 0.05$	$A(3)C(1)$	0.082	0.133	1.059	0.38
WSD^{dyn}_{Ag5}	$m_1 = 11, p = 0.15$	$A(1)C(1)$	0.122	0.226	1.136	2.56

The product rule

System	Parameters	Algorytm	e	e_{ONE}	$\overline{d}_{WSD^{dyn}_{Ag}}$	t
WSD^{dyn}_{Ag1}	$m_1 = 20, p = 0.15$	$A(3)C(3)$	0.085	0.085	1	0.06
WSD^{dyn}_{Ag2}	$m_1 = 16, p = 0.05$	$A(5)C(4)$	0.104	0.104	1	0.08
WSD^{dyn}_{Ag3}	$m_1 = 6, p = 0.05$	$A(1)C(2)$	0.122	0.234	1.157	0.07
WSD^{dyn}_{Ag4}	$m_1 = 1, p = 0.05$	$A(1)C(2)$	0.093	0.144	1.056	0.18
WSD^{dyn}_{Ag5}	$m_1 = 16, p = 0.05$	$A(1)C(1)$	0.138	0.237	1.144	2.33

The median rule

System	Parameters	Algorytm	e	e_{ONE}	$\overline{d}_{WSD^{dyn}_{Ag}}$	t
WSD^{dyn}_{Ag1}	$m_1 = 1, p = 0.15$	$A(3)C(9)$	0.109	0.109	1	0.07
	$m_1 = 6, p = 0.15$	$A(1)C(1)$	0.008	0.239	1.644	0.04
WSD^{dyn}_{Ag2}	$m_1 = 1, p = 0.15$	$A(2)C(2)$	0.189	0.202	1.016	0.06
	$m_1 = 11, p = 0.15$	$A(1)C(1)$	0.008	0.271	1.628	0.06
WSD^{dyn}_{Ag3}	$m_1 = 11, p = 0.15$	$A(2)C(3)$	0.117	0.160	1.069	0.11
	$m_1 = 11, p = 0.15$	$A(1)C(2)$	0.048	0.348	1.574	0.08
WSD^{dyn}_{Ag4}	$m_1 = 1, p = 0.15$	$A(7)C(2)$	0.250	0.279	1.045	1.31
	$m_1 = 1, p = 0.05$	$A(2)C(1)$	0.098	0.426	1.460	0.28
WSD^{dyn}_{Ag5}	$m_1 = 6, p = 0.15$	$A(3)C(5)$	0.348	0.418	1.112	3.01
	$m_1 = 1, p = 0.15$	$A(2)C(1)$	0.250	0.535	1.404	2.59

The maximum rule

System	Parameters	Algorytm	e	e_{ONE}	$\overline{d}_{WSD^{dyn}_{Ag}}$	t
WSD^{dyn}_{Ag1}	$m_1 = 1, p = 0.15$	$A(4)C(5)$	0.316	0.367	1.133	0.09
WSD^{dyn}_{Ag2}	$m_1 = 1, p = 0.15$	$A(2)C(5)$	0.449	0.601	2.064	0.06
WSD^{dyn}_{Ag3}	$m_1 = 1, p = 0.15$	$A(5)C(8)$	0.465	0.569	1.503	0.43
WSD^{dyn}_{Ag4}	$m_1 = 1, p = 0.15$	$A(4)C(5)$	0.590	0.737	1.521	0.49
WSD^{dyn}_{Ag5}	$m_1 = 1, p = 0.15$	$A(2)C(2)$	0.859	0.952	2.285	3.00

The minimum rule

System	Parameters	Algorytm	e	e_{ONE}	$\overline{d}_{WSD^{dyn}_{Ag}}$	t
WSD^{dyn}_{Ag1}	$m_1 = 1, p = 0.05$	$A(4)C(4)$	0.162	0.205	1.721	0.06
WSD^{dyn}_{Ag2}	$m_1 = 11, p = 0.15$	$A(2)C(6)$	0.176	0.245	1.846	0.07
WSD^{dyn}_{Ag3}	$m_1 = 6, p = 0.15$	$A(2)C(6)$	0.205	0.285	1.965	0.10
WSD^{dyn}_{Ag4}	$m_1 = 1, p = 0.05$	$A(6)C(1)$	0.096	0.229	1.846	1.29
WSD^{dyn}_{Ag5}	$m_1 = 1, p = 0.15$	$A(3)C(1)$	0.178	0.301	1.263	4.11

The results presented on the graph are divided into two groups. Methods that generate

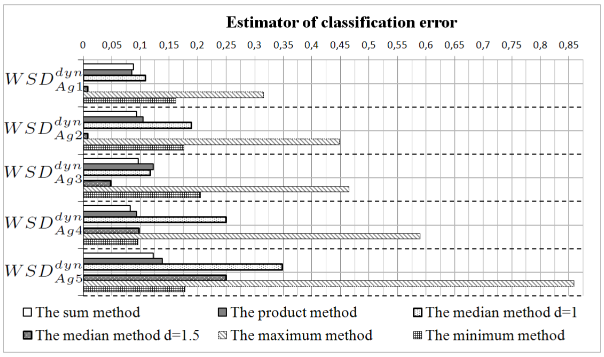

Fig. 1. Comparison of the results for the Soybean data set

a set of decisions with the average size approximately equal to 1: the sum, the product and the median rules $d = 1$, and the methods that generate larger set of global decisions: the median $d = 1.5$, the maximum and the minimum rules. Based on the results it can be concluded that among the methods that generate unambiguous decisions, the best method is the sum rule, in second place is the product and in third place is the median rule. When a system's user allows small ambiguity in the process of decisions making it is best to use the median rule. However, it should be noted that this method does well only in the case of a smaller number of resource agents. For a large number of agents 9 and 11 it does not generate good results. The minimum, and especially the maximum rule do not achieve satisfactory efficiency of inference, despite of significantly increased the average size of the global decisions sets. The results of the experiments with the Vehicle data set are presented in Table 2. As can be seen, for the Vehicle Silhouettes data set, unequivocal decisions are generated by all analyzed methods of combining classifiers' predictions. Which means that the average size of the global decisions sets is equal to 1. On the basis of detailed analysis of vectors of probabilities generated by the individual classifiers, it was concluded that the reason of this situation is that for the Vehicle data set there are no dummy agents. That is there are no undecided agents who assign the same probability value to many different decision values. Figure 2 shows a graphical comparison of the estimator of classification error for different dispersed systems of the Vehicle Silhouettes data set. Based on the presented results it can be concluded that the sum, the product and the median rules are significantly better than the maximum and the minimum rules. It is difficult to say which method is the best. For the systems with 3 and 5 resource agents (WSD^{dyn}_{Ag1}, WSD^{dyn}_{Ag2}) the best is the sum rule, in second place is the product and in third place is the median rule. But for the systems with 7 and 9 resource agents (WSD^{dyn}_{Ag3}, WSD^{dyn}_{Ag4}), the best is the median rule, then the product and the sum rule. For the system with 11 resource agents (WSD^{dyn}_{Ag5}) these three methods give the same result. Summarizing the results presented in tables 1 and 2 it can be said that the best results from the examined methods achieve the sum

Table 2. Summary of experiments results with the Vehicle Silhouettes data set

System	Parameters	Algorytm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}^{dyn}}$	t
The sum rule						
WSD_{Ag1}^{dyn}	$m_1 = 6, p = 0.05$	$A(3)C(4)$	0.240	0.240	1	0.08
WSD_{Ag2}^{dyn}	$m_1 = 1, p = 0.05$	$A(4)C(9)$	0.291	0.291	1	0.10
WSD_{Ag3}^{dyn}	$m_1 = 11, p = 0.05$	$A(4)C(6)$	0.252	0.252	1	0.20
WSD_{Ag4}^{dyn}	$m_1 = 6, p = 0.05$	$A(4)C(5)$	0.311	0.311	1	0.34
WSD_{Ag5}^{dyn}	$m_1 = 11, p = 0.05$	$A(4)C(1)$	0.268	0.268	1	2.51
The product rule						
WSD_{Ag1}^{dyn}	$m_1 = 6, p = 0.05$	$A(3)C(4)$	0.244	0.244	1	0.08
WSD_{Ag2}^{dyn}	$m_1 = 1, p = 0.05$	$A(8)C(13)$	0.295	0.295	1	0.16
WSD_{Ag3}^{dyn}	$m_1 = 11, p = 0.05$	$A(4)C(6)$	0.248	0.248	1	0.21
WSD_{Ag4}^{dyn}	$m_1 = 11, p = 0.05$	$A(5)C(8)$	0.303	0.303	1	1.16
WSD_{Ag5}^{dyn}	$m_1 = 11, p = 0.05$	$A(4)C(5)$	0.268	0.268	1	2.53
The median rule						
WSD_{Ag1}^{dyn}	$m_1 = 6, p = 0.05$	$A(7)C(6)$	0.252	0.252	1	0.10
WSD_{Ag2}^{dyn}	$m_1 = 1, p = 0.05$	$A(10)C(1)$	0.303	0.303	1.004	0.20
WSD_{Ag3}^{dyn}	$m_1 = 1, p = 0.05$	$A(4)C(8)$	0.240	0.240	1	0.13
WSD_{Ag4}^{dyn}	$m_1 = 1, p = 0.05$	$A(4)C(3)$	0.283	0.283	1	0.24
WSD_{Ag5}^{dyn}	$m_1 = 6, p = 0.05$	$A(4)C(5)$	0.268	0.272	1.004	2.48
The maximum rule						
WSD_{Ag1}^{dyn}	$m_1 = 1, p = 0.05$	$A(4)C(7)$	0.252	0.252	1	0.08
WSD_{Ag2}^{dyn}	$m_1 = 6, p = 0.05$	$A(5)C(4)$	0.327	0.327	1	0.14
WSD_{Ag3}^{dyn}	$m_1 = 1, p = 0.05$	$A(4)C(4)$	0.339	0.339	1	0.12
WSD_{Ag4}^{dyn}	$m_1 = 1, p = 0.05$	$A(5)C(5)$	0.362	0.366	1.004	0.32
WSD_{Ag5}^{dyn}	$m_1 = 11, p = 0.05$	$A(2)C(3)$	0.358	0.394	1.039	2.16
The minimum rule						
WSD_{Ag1}^{dyn}	$m_1 = 6, p = 0.05$	$A(3)C(2)$	0.287	0.287	1	0.08
WSD_{Ag2}^{dyn}	$m_1 = 1, p = 0.05$	$A(2)C(7)$	0.382	0.382	1	0.08
WSD_{Ag3}^{dyn}	$m_1 = 11, p = 0.05$	$A(4)C(2)$	0.280	0.283	1.004	0.20
WSD_{Ag4}^{dyn}	$m_1 = 11, p = 0.05$	$A(1)C(1)$	0.378	0.382	1.008	0.23
WSD_{Ag5}^{dyn}	$m_1 = 1, p = 0.05$	$A(5)C(1)$	0.343	0.366	1.024	2.34

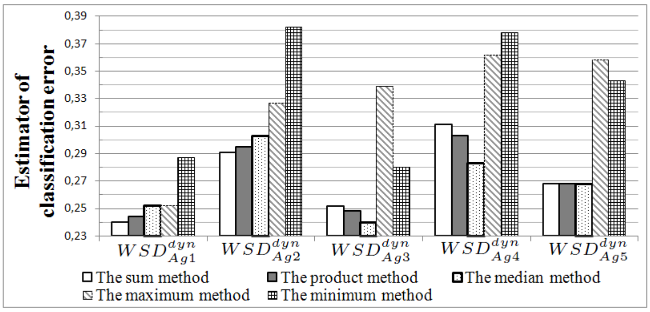

Fig. 2. Comparison of the results for the Vehicle Silhouettes data set

and product rule. These methods produce unambiguous results with the best observed efficiency of inference. During the experiments, the methods of combining classifiers' predictions without the use of a system with dynamically generated clusters were analyzed. The global decision-making process was as follows. On the basis of decision tables of resource agents the probability vectors were generated, and then one of the five discussed methods of combining predictions was used. The obtained results show that the use of a system with dynamically generated clusters significantly improves the efficiency of inference. However, due to the limited length of the article, results of these experiments are not presented here.

5 Conclusion

In this article, five different methods of conflict analysis were used in the dispersed decision-making system: the sum rule, the product rule, the median rule, the maximum rule and the minimum rule. In the experiments, which are presented, dispersed data have been used: Soybean data set and Vehicle Silhouettes data set. Based on the presented results of experiments it can be concluded that the sum and the product rules produce the best results from the methods that were examined. The maximum and the minimum rules produce the worst results. Especially, the results are not interesting, when dummy agents are present in a dispersed data set. It appears that the methods of conflict analysis should be applied in different situations and it seems to be possible to use more than one approach in the same session. Initially, several methods could be used simultaneously to generate the sets of global decisions, then these sets can be merged in some way. This could be an option for improving the overall accuracy. It is planned to investigate such an approach in a future work.

References

1. Cabrerizo, F.J., Herrera-Viedma, E., Pedrycz, W.: A method based on PSO and granular computing of linguistic information to solve group decision making problems defined in heterogeneous contexts. Eur. J. Oper. Res., 230(3), 624–633 (2013)

2. Delimata, P., Suraj, Z.: Feature Selection Algorithm for Multiple Classifier Systems: A Hybrid Approach. Fund. Informaticae 85 (1-4), 97–110 (2008)
3. Gatnar, E.: Multiple-model approach to classification and regression. PWN, Warsaw (2008)
4. Greco, S., Matarazzo, B., Słowiński R.: Rough sets theory for multicriteria decision analysis. Eur. J. Oper. Res., 129(1), 1–47 (2001)
5. Kittler, J., Hatef, M., Duin, R.P.W., Matas, J.: On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(3), 226–239 (1998)
6. Kuncheva, L., Bezdek, J.C., Duin, R.P.W.: Decision templates for multiple classifier fusion: an experimental comparison. Pattern Recognition, 34(2), 299–314 (2001)
7. Kuncheva, L.: Combining pattern classifiers methods and algorithms. John Wiley & Sons (2004)
8. Pawlak, Z.: On conflicts. Int. J. of Man-Machine Studies 21, 127–134 (1984)
9. Pawlak, Z.: An Inquiry Anatomy of Conflicts. Journal of Inform. Sciences 109, 65–78 (1998)
10. Przybyla-Kasperek, M., Wakulicz-Deja, A.: Application of reduction of the set of conditional attributes in the process of global decision-making, Fund. Informaticae 122 (4), 327-355 (2013)
11. Przybyla-Kasperek, M., Wakulicz-Deja, A.: Global decision-making system with dynamically generated clusters, Inform. Sciences, 270, 172–191 (2014)
12. Przybyla-Kasperek, M., Wakulicz-Deja, A.: A dispersed decision-making system - The use of negotiations during the dynamic generation of a systems structure. Inform. Sciences, 288, 194–219 (2014)
13. Schneeweiss, C.: Distributed decision making. Springer, Berlin (2003)
14. Schneeweiss, C.: Distributed decision making - a unified approach. Eur. J. Oper. Res., 150(2), 237–252 (2003)
15. Skowron, A., Wang, H., Wojna, A., Bazan, J.: Multimodal Classification: Case Studies. T. Rough Sets, 224–239 (2006)
16. Suraj, Z., Gayar, N.E., Delimata, P.: A Rough Set Approach to Multiple Classifier Systems. Fund. Informaticae 72 (1-3), IOS Press, Amsterdam, 393–406 (2006)
17. Wakulicz-Deja, A., PrzybyÅĆa-Kasperek, M.: Application of the method of editing and condensing in the process of global decision-making, Fund. Informaticae 106 (1), 93-117 (2011)
18. Wróblewski, J.: Ensembles of Classifiers Based on Approximate Reducts. Fund. Informaticae, 47(3-4), 351-360 (2001)
19. http://www.ics.uci.edu/ mlearn/MLRepository.html.