

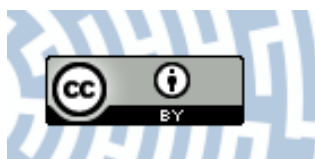


You have downloaded a document from  
**RE-BUŚ**  
repository of the University of Silesia in Katowice

**Title:** Coalitions' Weights in a Dispersed System with Pawlak Conflict Model

**Author:** Małgorzata Przybyła-Kasperek

**Citation style:** Przybyła-Kasperek Małgorzata. (2020). Coalitions' Weights in a Dispersed System with Pawlak Conflict Model. "Group Decision and Negotiation" iss. 3 (2020), s. 549-591, doi 10.1007/s10726-020-09667-1



Uznanie autorstwa - Licencja ta pozwala na kopiowanie, zmienianie, rozprowadzanie, przedstawianie i wykonywanie utworu jedynie pod warunkiem oznaczenia autorstwa.



UNIwersYTET ŚLĄSKI  
W KATOWICACH



Biblioteka  
Uniwersytetu Śląskiego



Ministerstwo Nauki  
i Szkolnictwa Wyższego



# Coalitions' Weights in a Dispersed System with Pawlak Conflict Model

Małgorzata Przybyła-Kasperek<sup>1</sup>

© The Author(s) 2020

## Abstract

The article addresses the issues related to making decisions by an ensemble of classifiers. Classifiers are built based on local tables, the set of local tables is called a dispersed knowledge. The paper discusses a novel application of Pawlak analysis model to examine the relations between classifiers and to create coalitions of classifiers. Each coalition has access to some aggregated knowledge on the basis of which joint decisions are made. Various types of coalitions are formed—a strong coalitions consisting of a large number and significant classifiers, and a weak coalitions consisting of insignificant classifiers. The new contributions of the paper is a systematic investigation of the weights of coalitions that influence the final decision. Four different method of calculating the strength of the coalitions have been applied. Each of these methods consider another aspect of the structure of the coalitions. Generally, it has been experimentally confirmed that, for a method that correctly identifies the relations between base classifiers, the use of coalitions weights improves the quality of classification. More specifically, it has been statistically confirmed that the best results are generated by the weighting method that is based on the size of the coalitions and the method based on the unambiguous of the decisions.

**Keywords** Conflict analysis · Dispersed knowledge · Multiple classifiers · Coalitions

## 1 Introduction

An important problem in today's world is the dispersion of knowledge. Many units, dealing with the same subject and operating in the same field, gather the knowledge to which they have access. This knowledge can be the result of various factors—experience, history, analyzed cases, sensors. Very popular form of saving knowledge is a decision table. However, if the knowledge contained in local decision tables is the result of different stimuli or analysis, the form of the tables can be very different,

---

✉ Małgorzata Przybyła-Kasperek  
malgorzata.przybyla-kasperek@us.edu.pl

<sup>1</sup> Institute of Computer Science, University of Silesia, Będzińska 39, 41-200 Sosnowiec, Poland

both in terms of the sets of conditional attributes and the sets of the universe. It is not possible to aggregate such knowledge. In this situation, a more sophisticated approach should be used.

In this paper the novel application of Pawlak rough set analysis to a dispersed system is considered. This approach was proposed in Przybyła-Kasperek (2017), and Przybyła-Kasperek (2017) for the first time, in this article it is further developed and analyzed.

For the simultaneous use of dispersed knowledge, a dispersed system with a two-stage aggregation structure was proposed. In the first step, local tables on the basis of which classifiers make similar decisions for the test object, are aggregated. In the second step, probability vectors generated based on aggregated tables are merged. In the paper (Przybyła-Kasperek 2017), Pawlak's conflict model was used to identify relations between base classifiers that are constructed based on local tables. In this way coalitions are generated for which the aggregated tables are defined.

The Pawlak's model was not originally designed for ensemble of classifiers. Therefore some modifications were proposed in the paper (Przybyła-Kasperek 2017) in order to apply the model to the problem of classification based on dispersed knowledge. Three approaches to adopting the model for the multiple classifier have been proposed. In the paper (Przybyła-Kasperek 2017), the structure of clusters created using these methods was analyzed. It turned out that the quality of the generated coalitions did not translate into the quality of the classification. The reason for this lies in the final decision-making method, where each coalition (both strong and weak) has the same influence on the decision.

The novelty that is proposed in this paper is to apply the weights of the coalitions in order to diversify the coalitions' influence on final decision. Four different methods of counting weights are analyzed. Coalitions' strength is calculated according to various factors. The first method takes into account the number of classifiers belonging to the coalition. The second method takes decisiveness of classifiers into account. The third method is a combination of the above two. The fourth method calculates the strength of the coalition depending on the variability of the vectors, which are generated based on the aggregated tables.

In this paper, these four methods have been applied to three approaches of using Pawlak's model in the ensemble of classifiers. It was shown on the example that the use of different approaches to determining weights of coalitions generates completely different results. In this paper in-depth and systematical experiments were presented that have been conducted on fifteen different dispersed sets (three different data sets, each dispersed in five different versions). Three different approaches to using Pawlak analysis model and four different methods for determining weights were compared. Statistical tests were performed to confirm that the use of coalitions' weights improved the quality of classification. It was found that the best results are generated by the weighting method that is based on the size of the coalitions and the method that is based on the unambiguous of the decisions.

To summarize the main contributions of this paper are as follows:

- proposition of four methods for determining the weights of coalitions in a dispersed system using Pawlak analysis model,

- examination of these methods for three approaches of application Pawlak analysis method in a dispersed system,
- comparison and statistical analysis of the obtained results.

In this paper, the problem of ensemble of classifiers is considered. However, the fundamental difference between the approaches known from the literature (Bloch 1996; Shoemaker et al. 2008; Stefanowski 2005; Tang et al. 2006) and the one considered in the paper lies in the knowledge representation to which the classifiers have access. In the paper it was assumed that knowledge is pre-specified and accumulated separately by independent units. Therefore, we can not expect that the sets of objects or the sets of attributes fulfill some relations (inclusion, equality, or disjunction). Therefore, the approach that is considered in the paper is more general. In Panov and Džeroski (2007), a classifier ensembles with different sets of objects and different sets of attributes are considered. This method is a combination of Bagging and Random Subspace Method. However, the knowledge that is provided in this model is stored in a single data base and it is not possible to use several local decision tables that were collected separately. In Polikar et al. (2006), a system with incremental learning capability was proposed. This method allows to learn new information when a new data set is available. Data sets have different sets of attributes. This approach consists in creating several classifiers based on each new set of data. The system does not analyze the relationship between the available data sets and does not check the consistency of the knowledge contained in them.

In the paper the approach is used in which the recognition of relations and the creation of coalitions are very important. In the literature, different approaches to this issue can be found (Lopes et al. 2008; Fatima et al. 2005; Kersten and Lai 2007; Rahwan et al. 2004). In this article the Pawlak's model (Pawlak 1984) is used. The model allows to make advanced analysis of relations between agents and is a simple way to illustrate the basic properties of conflicts. The Pawlak's model has been studied by many authors (Deja 2002; Ramanna et al. 2006; Ramanna et al. 2007; Skowron and Deja 2002; Skowron et al. 2006). In Lang et al. (2017) a probabilistic model of conflict analysis was proposed, which is a combination of the Pawlak model and three-way decisions approach. Instead of using one threshold value, as it is done in this paper, two thresholds were used in order to recognize relations in conflict situation. In Yao (2019) three levels of conflict (strong conflict, weak conflict, and non-conflict) were considered. The author noticed some inconsistencies in the Pawlak model and removed them. In this study a completely different application of the Pawlak's model is considered.

The article is organized as follows. In the second section the three methods of using Pawlak's model for ensemble of classifiers are discussed. The third section describes the general way of operating a dispersed system. In the fourth section, methods of calculating weights are presented. The fifth section contains the example of applying the coalitions' weights in a dispersed system. The sixth section shows the experimental protocol. In the seventh section the results of experiments carried out using some data sets from the UCI repository are presented. The article concludes with a short summary.

## 2 Application of the Pawlak's Model for Ensemble of Classifiers

The Pawlak's original model was proposed (Pawlak 1984, 2005, 2006) to analyze the conflict situation in which the agents (parties participating in the conflict) decided to resolve the dispute peacefully. In the model, each agent express his views on some issues, from the set  $A$ , by assigning one of three values. For each issue, each agent assigns a value: 1—means the agent is favoring the issue or 0—means the agent is neutral or  $-1$ —means the agent is opposed to the issue. Such information about the conflict situation is stored in the form of an information system  $S = (U, A)$ , where the universe  $U$  are agents,  $A$  is a set of issues, and the set of values of  $a \in A$  is equal  $V^a = \{-1, 0, 1\}$ . Opinion of agent  $ag$  about issue  $a$  is the value  $a(ag)$ .

In the Pawlak's model, based on such information system, relations between agents are defined and coalitions of agents in friendship (clique) are created.

In the paper (Przybyła-Kasperek 2017), this model was used to the problem of ensemble of classifiers. It is assumed that each classifier is interpreted as an agent in the Pawlak's model. The classifier  $ag$  generates a vector of ranks  $\mathbf{r}_{ag}(x) = [r_{ag,1}(x), \dots, r_{ag,c}(x)]$  that reflects the classification of the test object  $x$  ( $c$  is the number of decision classes). This is implemented as follows. At first, a vector of values  $\bar{\mu}_{ag}(x) = [\bar{\mu}_{ag,1}(x), \dots, \bar{\mu}_{ag,c}(x)]$  is generated based on each local decision table using the modified  $m_1$ -Nearest Neighbor algorithm. That is, from each local decision table and each decision class,  $m_1$  objects that are the most similar to the test object are selected. The vector's coordinate for a given decision value is equal to the average similarity of the relevant objects from that decision class. Based on this vector, a vector of ranks  $\mathbf{r}_{ag}(x) = [r_{ag,1}(x), \dots, r_{ag,c}(x)]$  for classifier  $ag$  is generated. For example, if we have five decision classes and the vector  $\bar{\mu}_{ag}(x) = [0.8, 0.4, 0.6, 0.4, 1]$  then the vector of ranks is equal to  $\mathbf{r}_{ag}(x) = [2, 4, 3, 4, 1]$ . Which means that the decision with the highest value in the vector  $\bar{\mu}_{ag}(x)$  has the rank 1, the second decision in the order has the rank 2, and so on.

A set of issues  $A$  that is considered in the Pawlak's conflict model is a set of decision attribute values. Based on the vectors of ranks, the views of the classifiers on the set of issues  $A$  are defined. Two approaches have been proposed to generate an information system.

In both methods, the information system  $S = (U, A)$ , where  $U$  is a set of classifiers,  $A$  is a set of decision attribute values, is defined. In the first method the vectors of ranks are converted into the views of the classifiers in the following way. Function  $a : U \rightarrow \{-1, 0, 1\}$  for each  $a \in A$  is defined

$$a(ag) = \begin{cases} 1 & \text{if } r_{ag,a}(x) = 1 \\ -1 & \text{if } r_{ag,a}(x) > 1 \end{cases} \quad (1)$$

In the second method it is realized in the following way. Function  $a : U \rightarrow \{-1, 0, 1\}$  for each  $a \in A$  is defined

$$a(ag) = \begin{cases} 1 & \text{if } r_{ag,a}(x) = 1 \\ 0 & \text{if } r_{ag,a}(x) = 2 \\ -1 & \text{if } r_{ag,a}(x) > 2 \end{cases} \tag{2}$$

As can be seen, in the first method the classifier is favoring the decision only if it has the rank 1; neutrality was not used here. In the second method the classifier is favorable to the decisions with the rank 1 and it is neutral to the decisions with the rank 2. It may be concluded that, the first method of defining an information system is more restrictive.

In the Pawlak’s model, the relations between classifiers are defined based on the information system in one of the two following ways. A function of distance between agents can be used. The function  $\rho_B^* : U \times U \rightarrow [0, 1]$  for the set of issues  $B \subseteq A$  is defined

$$\rho_B^*(ag_1, ag_2) = \frac{\sum_{a \in B} \phi_a^*(ag_1, ag_2)}{\text{card}\{B\}}, \tag{3}$$

where

$$\phi_a^*(ag_1, ag_2) = \begin{cases} 0 & \text{if } a(ag_1)a(ag_2) = 1 \text{ or } ag_1 = ag_2, \\ 0.5 & \text{if } a(ag_1)a(ag_2) = 0 \text{ and } ag_1 \neq ag_2, \\ 1 & \text{if } a(ag_1)a(ag_2) = -1. \end{cases}$$

Or a conflict function can be used. The function  $\rho_B : U \times U \rightarrow [0, 1]$  for the set of issues  $B \subseteq A$  is defined as follows

$$\rho_B(ag_1, ag_2) = \frac{\text{card}\{\delta_B(ag_1, ag_2)\}}{\text{card}\{B\}}, \tag{4}$$

where  $\delta_B(ag_1, ag_2) = \{a \in B : a(ag_1) \neq a(ag_2)\}$ .

The main difference between these functions refers to neutral classifiers. We will explain this by example. Let it be given the issue  $a \in A$  and the three classifiers  $ag_1, ag_2, ag_3 \in U$ . We assume that  $a(ag_1) = 1, a(ag_2) = 0, a(ag_3) = -1$ . Then  $\rho_{\{a\}}^*(ag_1, ag_2) = 0, 5$  and  $\rho_{\{a\}}^*(ag_1, ag_3) = 1$ . Thus, the distance of classifiers who are neutral is less than the distance of classifiers who are in conflict. For the conflict function, we have  $\rho_{\{a\}}(ag_1, ag_2) = 1$  and  $\rho_{\{a\}}(ag_1, ag_3) = 1$ . Thus, the distance between neutral classifiers and classifiers in conflict is the same. This can be summarized that the conflict function is more restrictive.

In the Pawlak’s model, a pair of classifiers  $ag_1, ag_2 \in U$  is said to be:

- allied  $R^+(x, y)$ , if  $\rho^*(ag_1, ag_2) < 0.5$  (or  $\rho(ag_1, ag_2) < 0.5$ ),
- in conflict  $R^-(x, y)$ , if  $\rho^*(ag_1, ag_2) > 0.5$  (or  $\rho(ag_1, ag_2) > 0.5$ ),
- neutral  $R^0(x, y)$ , if  $\rho^*(ag_1, ag_2) = 0.5$  (or  $\rho(ag_1, ag_2) = 0.5$ ).

Set  $X \subseteq U$  is a coalition if for every  $ag_1, ag_2 \in X, R^+(ag_1, ag_2)$  and  $ag_1 \neq ag_2$ .

In the first method of defining the information system, there is no 0 value, so there are no neutral classifiers. This means that both functions assign the same values for each pair of classifiers. In the second method of defining the information system, neutral classifiers occur, so functions can have different values. As was explained above, the distance function is less restrictive for neutral agents—it assigns smaller value for neutral agents—than the conflict function. Based on the value of one of the functions, we define a coalition as a subset of classifiers such that for each pair of classifiers the value of the function is less than 0.5. Finally, three different approaches to generating coalitions were considered:

- *Approach 1*—the first method of defining an information system is used,
- *Approach 2*—the second method of defining an information system and the function of distance between agents are used,
- *Approach 3*—the second method of defining an information system and the conflict function are used.

In the next section, issues related to the use of the created coalitions and the organization of the dispersed system, are described.

### 3 Scheme of a Dispersed System's Operation

Methods of creating coalitions that were described in the previous section can be applied to any ensemble of classifiers. In order to perform the experiments, some technical specifications have been adopted in this paper. These are described below. The way to generate vectors of ranks and the method for aggregation local decision tables have been taken from the previous work of the author (Przybyła-Kasperek 2017; Przybyła-Kasperek and Wakulicz-Deja 2016).

General rules of operation of a dispersed system can be described in several steps. In the first step, the coalitions of classifiers are determined using one of the three approaches discussed in the previous section.

In the second step, on the basis of local decision tables from one coalition an aggregated decision table is generated. The method of approximation aggregation is described in the paper (Przybyła-Kasperek and Wakulicz-Deja 2016). The method is implemented as follows. From each of the local tables and from each decision class,  $m_2$  objects that are the most similar to the test object are selected. Then the objects are merged under certain conditions. If objects come from the same decision class and have consistent values on common attributes, then they are written as one object in the aggregated table.

In the third step, a vector of values is determined on the basis of each aggregated table. The vector's coordinate for a given decision value is equal to the maximum similarity of objects from the decision class and the aggregated table to the test object.

In the last step, some linear transformations are made on these vectors of values and global decisions is generated. However, this will be discussed in the next section, as the novelty—the coalitions' weights—will be applied at this stage.

## 4 Weights for Coalitions

The methods of creating coalitions of classifiers and identifying relations between classifiers that were described in Sect. 2, have already been analyzed in the paper (Przybyła-Kasperek 2017). In this article the structures of created coalitions were compared in detail. It was found that approaches 1 and 2 tend to create a smaller number of large coalitions. For data sets with a large number of decision classes, the extreme situation was obtained, in which practically one coalition consisting of all classifiers was generated. This is a very unfavorable situation and means that, in this case, these approaches have lost the ability to identify the relations between the classifiers. Only approach 3, which uses the second method of defining an information system and the conflict function, generates different result. A larger number of smaller coalitions has been obtained in this case. In addition, in this approach, classifiers are more likely to be simultaneously assigned to several coalitions. Finally, it was concluded that approach 3 has the greatest ability to identify relations between classifiers.

In the second step of the experiments that were presented in article (Przybyła-Kasperek 2017), the quality of the classification was analyzed using these three approaches. It was found that significant differences in coalitions' structure do not translate into the quality of classification. The reason for this lies in the method that was used to generate final decisions. Some novelties are proposed in this paper so that the coalitions' structure is taken into account when generating the final result. The method that was used previously and the introduced modifications are described below.

As was mentioned earlier, after determination of the classifiers' coalitions, for each coalition  $C_j$  an aggregated decision table is generated. Based on each of the aggregated tables, the vector of values  $\mu_j(x) = [\mu_{j,1}(x), \dots, \mu_{j,c}(x)]$ , where  $c$  is the number of decision classes and  $x$  is the classified object, is generated as described in Sect. 3. Each vector coordinate corresponds to one decision value. The final decisions are generated using these vectors.

In the previous papers, the method of density-based algorithm was used. It was implemented as follows. First, the sum of the vectors  $\mu_j(x)$  was calculated. Then a decision that has the maximum value of the vector's coordinates  $\sum_j \mu_j(x)$  is selected. Finally, the DBSCAN algorithm was used in order to determine which values of the decision have vector's coordinates densely located around the maximum value (they are close enough to the maximum value). In this way, not only the decision which has the maximum support of all classifiers but also the ones for which support is sufficiently large is determined. However, if a simple sum of vectors is considered, each coalition has the same influence on the final decision. The coalition's structure and its size are not taken into account. In this situation, the process of coalitions' generation, and better identification of relations between classifiers is less important, and is not fully exploited.

Therefore, in this work, instead of a simple sum of the vectors, the weighted sum will be calculated. The weight is assigned to each coalition, which takes into



account the size, the structure and the commitment of the members into the coalition formation. Four different methods of calculating weights are considered.

The first method—the size of the coalition. In this method, a weight that takes into account the number of classifiers belonging to the coalition is assigned to each coalition. In all three approaches of applying the Pawlak’s model, inseparable coalitions are generated. And that is why it is important to take into account the partial involvement of the classifier in the formation of the coalition. For each classifier  $ag$  the coefficient  $m_{ag}^x$ , where  $x$  is the classified object, is determined. The value of this coefficient is inversely proportional to the number of coalitions to which the classifier belongs

$$m_{ag}^x = (\text{card}\{C_j^x : ag \in C_j^x\})^{-1}, \tag{5}$$

where  $C_1^x, \dots, C_m^x$  are the coalitions for object  $x$ . The weight that is assigned to the coalition  $C_j^x$  is equal to  $\sum_{ag \in C_j^x} m_{ag}^x$ . In this way, the larger coalitions will have a greater influence on the final decisions that were made.

The second method—the unambiguous of the decisions made by the classifiers. The quality of the decisions taken by classifiers belonged to the coalition should also be taken into account when determining the weight of the coalition. By the quality of the decisions we understand the uniqueness of decisions, in other words decisiveness of classifiers. If the classifier  $ag$  generates a vector of values  $\bar{\mu}_{ag}(x)$  (defined in Sect. 3) in which the coordinates are very different, i.e. one of them is significantly larger than the others, it means that the classifier is certain of the taken decision. The probability of classifying the test object into one of the classes is much higher compared to the probabilities for other classes. The standard deviation of the values of the vector’s coordinates is the measure of the quality of the taken decisions that correspond to such assumptions. Thus, for each classifier  $ag$  the standard deviation of the values of the vector  $\bar{\mu}_{ag}(x)$  is calculated, denoted as  $sd_{ag}^x$ . The weight that is assigned to the coalition  $C_j^x$  is equal to

$$\frac{\sum_{ag \in C_j^x} sd_{ag}^x}{\sum_{j=1}^m \sum_{ag' \in C_j^x} sd_{ag'}^x}. \tag{6}$$

In this way, coalitions that contain more resolute and assertive classifiers will have a greater influence on the final decisions that were made.

The third method—the size of the coalition and the unambiguous of the taken decisions. In this method two previously discussed approaches have been combined. It means that both the size of the coalition expressed by the number of classifiers included in it and the assertiveness of these classifiers are taken into account. For the classifier, the two defined before coefficients (the partial involvement of the classifier in the

**Table 1** Test object  $x$

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$a_{15}$
$x$	1	1	0	2	1	0	2	2	1	2	2	2	1	2	1

Table 2 Relevant objects from the decision tables

$U_{ag1}$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	d	$s(x_i, x)$
$x_1^1$	1	1	1	2	1	$v_1$	1
$x_2^1$	2	2	2	2	1	$v_2$	2
$x_3^1$	1	2	0	2	2	$v_3$	2
$x_4^1$	0	0	0	1	1	$v_4$	1
$x_5^1$	1	1	0	2	1	$v_5$	1

$U_{ag3}$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	d	$s(x_i, x)$
$x_1^3$	1	0	0	2	0	$v_1$	1
$x_2^3$	2	1	2	0	2	$v_2$	2
$x_3^3$	2	2	1	2	1	$v_3$	2
$x_4^3$	0	2	1	0	2	$v_4$	1
$x_5^3$	2	2	1	2	2	$v_5$	1

$U_{ag4}$	$a_1$	$a_2$	$a_{12}$	d	$s(x_i, x)$
$x_1^4$	2	1	2	$v_1$	1
$x_2^4$	1	2	2	$v_2$	2
$x_3^4$	1	0	0	$v_3$	2
$x_4^4$	0	0	2	$v_4$	1
$x_5^4$	0	1	1	$v_5$	2

$U_{ag5}$	$a_3$	$a_{13}$	$a_{14}$	d	$s(x_i, x)$
$x_1^5$	0	2	2	$v_1$	1
$x_2^5$	1	1	2	$v_2$	2
$x_3^5$	2	1	1	$v_3$	1
$x_4^5$	0	1	0	$v_4$	2
$x_5^5$	0	1	2	$v_5$	1

$U_{ag6}$	$a_6$	$a_{14}$	$a_{15}$	d	$s(x_i, x)$
$x_1^6$	0	2	1	$v_1$	1
$x_2^6$	1	1	1	$v_2$	2
$x_3^6$	0	2	0	$v_3$	2
$x_4^6$	0	2	2	$v_4$	1
$x_5^6$	2	2	1	$v_5$	2

$U_{ag7}$	$a_4$	$a_5$	$a_{10}$	$a_{11}$	$a_{12}$	d	$s(x_i, x)$
$x_1^7$	0	1	2	2	2	$v_1$	1
$x_2^7$	1	1	1	1	2	$v_2$	2
$x_3^7$	2	0	0	2	2	$v_3$	2
$x_4^7$	2	1	2	2	0	$v_4$	1
$x_5^7$	2	1	2	2	1	$v_5$	2

$U_{ag2}$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	d	$s(x_i, x)$
$x_1^2$	1	0	0	2	2	$v_1$	1
$x_2^2$	2	0	1	2	2	$v_2$	2
$x_3^2$	2	2	2	2	2	$v_3$	2
$x_4^2$	2	1	0	1	1	$v_4$	1
$x_5^2$	2	1	1	1	1	$v_5$	1

**Table 3** Vectors  $\bar{\mu}_{ag}(x)$  and  $\mathbf{r}_{ag}(x)$

Classifier	$\bar{\mu}_{ag}(x)$	$\mathbf{r}_{ag}(x)$	$sd_{ag}^x$ —standard deviation of normalized $\bar{\mu}_{ag}(x)$
$ag_1$	[0.8, 0.4, 0.6, 0.4, 1]	[2,4,3,4,1]	0.166
$ag_2$	[0.6, 0.6, 0.6, 0.6, 0.4]	[1,1,1,1,2]	0.065
$ag_3$	[0.2, 0.4, 0.8, 0.6, 1]	[5,4,2,3,1]	0.215
$ag_4$	[0.67, 0.67, 0.33, 0.33, 0.33]	[1,1,2,2,2]	0.159
$ag_5$	[0.67, 0.67, 0.33, 0.67, 1]	[2,2,3,2,1]	0.144
$ag_6$	[1, 0.33, 0.67, 0.67, 0.67]	[1,3,2,2,2]	0.144
$ag_7$	[0.8, 0.4, 0.6, 0.8, 0.8]	[1,3,2,1,1]	0.107

**Table 4** Information systems

Formula 1					
U	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
$ag_1$	-1	-1	-1	-1	+1
$ag_2$	+1	+1	+1	+1	-1
$ag_3$	-1	-1	-1	-1	+1
$ag_4$	+1	+1	-1	-1	-1
$ag_5$	-1	-1	-1	-1	+1
$ag_6$	+1	-1	-1	-1	-1
$ag_7$	+1	-1	-1	+1	+1

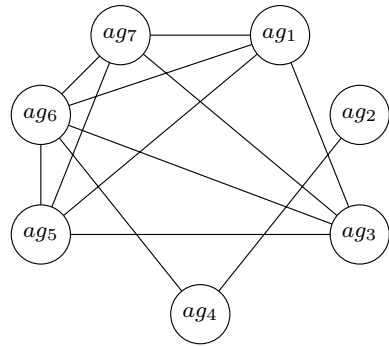
  

Formula 2					
U	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$
$ag_1$	0	-1	-1	-1	+1
$ag_2$	+1	+1	+1	+1	0
$ag_3$	-1	-1	0	-1	+1
$ag_4$	+1	+1	0	0	0
$ag_5$	0	0	-1	0	+1
$ag_6$	+1	-1	0	0	0
$ag_7$	+1	-1	0	+1	+1

**Table 5** Function values—  
Approach 1

	$ag_1$	$ag_2$	$ag_3$	$ag_4$	$ag_5$	$ag_6$	$ag_7$
$ag_1$							
$ag_2$	1						
$ag_3$	0	1					
$ag_4$	0.6	0.4	0.6				
$ag_5$	0	1	0	0.6			
$ag_6$	0.4	0.6	0.4	0.2	0.4		
$ag_7$	0.4	0.6	0.4	0.6	0.4	0.4	

**Fig. 1** A graphical representation—Approach 1



formation of the coalition and the standard deviation of the generated vector) are multiplied. The weight for the coalition  $C_j^x$  is calculated as follows

$$\sum_{ag \in C_j^x} m_{ag}^x \cdot sd_{ag}^x. \tag{7}$$

The fourth method—the unambiguous of the decisions made by the coalition. At first glance this method looks the same as the second method. However, the basic difference is that this time when calculating the weight of the coalition, a vector that was generated based on the aggregated decision table of the coalition is considered. The value of the weight is not dependent on the decisions taken by the coalition's members, but on aggregate knowledge of the coalition. Analogously to the second method, the standard deviation  $SD_j^x$  of the vector  $\mu_j(x)$  is calculated for each coalition. The weight for the coalition  $C_j^x$  is calculated as follows

$$\frac{SD_j^x}{\sum_{k=1}^m SD_k^x}, \tag{8}$$

where  $C_1^x, \dots, C_m^x$  are the coalitions for object  $x$ .

The method of generating the final decision is as follows. A vector equal to the weighted average of the vectors generated based on the aggregated tables is calculated

$$\sum_j \omega_j^x \cdot \mu_j(x),$$

where  $\omega_j^x$  is the weight generated for the coalition  $C_j^x$  and the test object  $x$  according to one of the four approaches described above. Then, as before, a decision with a maximum vector's coordinates is chosen. Finally, the DBSCAN algorithm is used in order to select a set of decisions that are densely located around the decision with the maximum support of the classifiers.

**Table 6** Aggregated decision tables—Approach 1

Coalition $\{ag_1, ag_3, ag_5, ag_6, ag_7\}$																
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$a_{15}$	$d$	$s(x_i, x)$
?	?	0	0	1	0	?	?	?	2	2	2	2	2	1	$v_1$	0.800
?	?	0	?	?	0	1	0	0	2	0	?	2	2	1	$v_1$	0.500
1	1	1	2	1	0	1	0	0	2	0	?	?	2	1	$v_1$	0.615
?	?	?	1	1	1	?	?	?	1	1	2	?	1	1	$v_2$	0.375
?	?	1	1	1	?	?	?	?	1	1	2	1	2	?	$v_2$	0.500
2	2	2	2	1	1	2	1	2	0	2	?	?	1	1	$v_2$	0.385
?	?	1	?	?	?	2	1	2	0	2	?	1	2	?	$v_2$	0.500
?	?	?	2	0	0	?	?	?	0	2	2	?	2	0	$v_3$	0.625
?	?	2	2	0	?	?	?	?	0	2	2	1	1	?	$v_3$	0.500
?	?	2	?	?	?	2	2	1	2	1	?	1	1	?	$v_3$	0.692
1	2	0	2	2	0	2	2	1	2	1	?	?	2	0	$v_3$	0.625
?	?	?	2	1	0	?	?	?	2	2	0	?	2	2	$v_4$	0.750
?	?	0	2	1	?	?	?	?	2	2	0	1	0	?	$v_4$	0.750
0	0	0	1	1	?	0	2	1	0	2	?	1	0	?	$v_4$	0.538
0	0	0	1	1	0	0	2	1	0	2	?	?	2	2	$v_4$	0.500
1	1	0	2	1	2	2	2	1	2	2	1	1	2	1	$v_5$	0.867
Coalition $\{ag_2, ag_4\}$																
$a_1$	$a_2$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_{12}$	$d$	$s(x_i, x)$							
2	1	1	0	0	2	2	2	$v_1$	0.625							
1	2	2	0	1	2	2	2	$v_2$	0.625							
1	0	2	2	2	2	2	0	$v_3$	0.500							
0	0	2	1	0	1	1	2	$v_4$	0.500							

Table 6 (continued)

Coalition $\{ag_2, ag_4\}$										
$a_1$	$a_2$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_{12}$	$d$	$s(x_i, x)$	
0	1	2	1	1	1	1	1	$v_5$	0.375	
Coalition $\{ag_4, ag_6\}$										
$a_1$	$a_2$	$a_6$	$a_{12}$	$a_{14}$	$a_{15}$	$d$	$s(x_i, x)$			
2	1	0	2	2	1	$v_1$	0.833			
1	2	1	2	1	1	$v_2$	0.500			
1	0	0	0	2	0	$v_3$	0.500			
0	0	0	2	2	2	$v_4$	0.500			
0	1	2	1	2	1	$v_5$	0.500			

**Table 7** Vectors of coalitions—Approach 1

Coalition	Vector	Vector transformation	$SD_j^x$ — standard deviation
$\{ag_1, ag_3, ag_5, ag_6, ag_7\}$	[0.800,0.500,0.692,0.750,0.867]	[0.222,0.139,0.192,0.208,0.240]	0.035
$\{ag_2, ag_4\}$	[0.625,0.625,0.500,0.500,0.375]	[0.238,0.238,0.190,0.190,0.143]	0.036
$\{ag_4, ag_6\}$	[0.833,0.500,0.500,0.500,0.500]	[0.294,0.176,0.176,0.176,0.176]	0.047

**Table 8** Coalitions' weights and weighted average of the vectors—Approach 1

Coalition	Weight	Vector
First method—coefficients $m_{ag}^x$		
$\{ag_1, ag_3, ag_5, ag_6, ag_7\}$	$4.5/7 = 0.643$	[0.143,0.089,0.123,0.134,0.154]
$\{ag_2, ag_4\}$	$1.5/7 = 0.214$	[0.051,0.051,0.041,0.041,0.031]
$\{ag_4, ag_6\}$	$1/7 = 0.143$	[0.042,0.025,0.025,0.025,0.025]
Sum		[0.236,0.165,0.189,0.200,0.210]
Second method—coefficients $sd_{ag}^x$		
$\{ag_1, ag_3, ag_5, ag_6, ag_7\}$	0.595	[0.132,0.082,0.114,0.124,0.143]
$\{ag_2, ag_4\}$	0.172	[0.041,0.041,0.033,0.033,0.025]
$\{ag_4, ag_6\}$	0.233	[0.068,0.041,0.041,0.041,0.041]
Sum		[0.241,0.165,0.188,0.198,0.209]
Third method—coefficients $m_{ag}^x \cdot sd_{ag}^x$		
$\{ag_1, ag_3, ag_5, ag_6, ag_7\}$	0.704	[0.156,0.097,0.135,0.146,0.169]
$\{ag_2, ag_4\}$	0.145	[0.034,0.034,0.028,0.028,0.021]
$\{ag_4, ag_6\}$	0.152	[0.045,0.027,0.027,0.027,0.027]
Sum		[0.235,0.159,0.189,0.201,0.216]
Fourth method—coefficients $SD_j^x$		
$\{ag_1, ag_3, ag_5, ag_6, ag_7\}$	0.295	[0.065,0.041,0.057,0.061,0.071]
$\{ag_2, ag_4\}$	0.401	[0.072,0.072,0.058,0.058,0.043]
$\{ag_4, ag_6\}$	0.304	[0.118,0.071,0.071,0.071,0.071]
Sum		[0.256,0.184,0.185,0.190,0.185]

## 5 Example of Applying the Coalitions' Weights in a Dispersed System

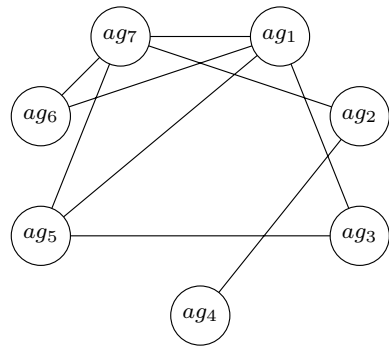
In this section, an example of the use of three approaches to creating coalitions of classifiers and four methods of calculating coalitions weights will be presented. The example will also illustrate the differences in the results generated by these methods.

We assume that we have the set of seven classifiers (interpreted as agents in the Pawlak's model)  $U = \{ag_1, ag_2, ag_3, ag_4, ag_5, ag_6, ag_7\}$  and the set of five decision values  $A = \{v_1, v_2, v_3, v_4, v_5\}$ . Each classifier is built based on one decision table  $D_{ag_i} = (U_{ag_i}, A_{ag_i}, d_{ag_i})$ ,  $i \in \{1, \dots, 7\}$ , in this example we assume that the conditional attributes in decision tables are qualitative. For the test object  $x$ , the classifier  $ag$ , generates a vector of ranks  $r_{ag}(x)$  by using the  $m_1$ -nearest neighbor algorithm. We

**Table 9** Function values—  
Approach 2

	$ag_1$	$ag_2$	$ag_3$	$ag_4$	$ag_5$	$ag_6$	$ag_7$
$ag_1$							
$ag_2$	0.8						
$ag_3$	0.2	0.8					
$ag_4$	0.6	0.3	0.7				
$ag_5$	0.3	0.6	0.4	0.5			
$ag_6$	0.4	0.5	0.5	0.5	0.5		
$ag_7$	0.4	0.4	0.5	0.5	0.4	0.3	

**Fig. 2** A graphical representation—Approach 2



assume that  $m_1 = 1$  and the test object is presented in Table 1. The nearest objects from each decision class of the decision tables and the values of the Gower similarity measure  $s(x, y)$  (that in the case of qualitative attributes is equivalent to the Hamming measure) are given in Table 2. Based on these similarity values, the vectors of values  $\bar{\mu}_{ag}(x) = [\bar{\mu}_{ag,1}(x), \dots, \bar{\mu}_{ag,c}(x)]$ ,  $ag \in U, c = card\{A\}$  are generated, and then the rank vectors  $\mathbf{r}_{ag}(x) = [r_{ag,1}(x), \dots, r_{ag,c}(x)]$  are determined. Both are given in Table 3. In order to determine the coalitions of classifiers, the information system  $S = (U, A)$  is defined. Depending on the approach (*Approach 1*, *Approach 2* or *Approach 2*), this is carried out in accordance with Formula 1 or Formula 2. Both information systems are presented in Table 4.

Based on the information systems, the values of the function of distance between agents (Formula 3) or the values of the conflict function (Formula 4) are determined, depending on the approach. From now on, the example will be solved separately for each approach.

### 5.1 Approach 1

In *Approach 1* we use the first method of defining an information system (Formula 1). In this case both functions (Formula 3 and Formula 4) are equivalent. The



**Table 10** Aggregated decision tables—Approach 2

Coalition $\{ag_1, ag_3, ag_5\}$														
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{13}$	$a_{14}$	$d$	$s(x_i, x)$	
?	?	0	?	?	1	0	0	2	0	2	2	$v_1$	0.375	
1	1	1	2	1	1	0	0	2	0	?	?	$v_1$	0.500	
?	?	1	?	?	2	1	2	0	2	1	2	$v_2$	0.500	
2	2	2	2	1	2	1	2	0	2	?	?	$v_2$	0.400	
?	?	2	?	?	2	2	1	2	1	1	1	$v_3$	0.625	
1	2	0	2	2	2	2	1	2	1	?	?	$v_3$	0.700	
0	0	0	1	1	0	2	1	0	2	1	0	$v_4$	0.500	
1	1	0	2	1	2	2	1	2	2	1	2	$v_5$	1	
Coalition $\{ag_1, ag_5, ag_7\}$														
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$s(x_i, x)$				
?	?	0	0	1	2	2	2	2	2	$v_1$	0.750			
1	1	1	2	1	?	?	?	?	?	$v_1$	0.800			
?	?	1	1	1	1	1	2	1	2	$v_2$	0.500			
2	2	2	2	1	?	?	?	?	?	$v_2$	0.400			
?	?	2	2	0	0	2	2	1	1	$v_3$	0.500			
1	2	0	2	2	?	?	?	?	?	$v_3$	0.600			
?	?	0	2	1	2	2	0	1	0	$v_4$	0.750			
0	0	0	1	1	?	?	?	1	0	$v_4$	0.429			
1	1	0	2	1	2	2	1	1	2	$v_5$	0.900			
Coalition $\{ag_1, ag_6, ag_7\}$														
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{14}$	$s(x_i, x)$				
?	?	?	0	1	0	2	2	2	2	$v_1$	0.875			

Table 10 (continued)

Coalition $\{ag_1, ag_6, ag_7\}$												
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{14}$	$a_{15}$	$d$	$s(x_i, x)$
1	1	1	2	1	0	?	?	?	2	1	$v_1$	0.875
?	?	?	1	1	1	1	1	2	1	1	$v_2$	0.375
2	2	2	2	1	1	?	?	?	1	1	$v_2$	0.375
?	?	?	2	0	0	0	2	2	2	0	$v_3$	0.625
1	2	0	2	2	0	?	?	?	2	0	$v_3$	0.625
0	0	0	1	1	0	?	?	?	2	2	$v_4$	0.750
?	?	?	2	1	0	2	2	0	2	2	$v_4$	0.500
1	1	0	2	1	2	2	2	1	2	1	$v_5$	0.818
Coalition $\{ag_2, ag_4\}$												
$a_1$	$a_2$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_{12}$	$d$	$s(x_i, x)$			
2	1	1	0	0	2	2	2	$v_1$	0.625			
1	2	2	0	1	2	2	2	$v_2$	0.625			
1	0	2	2	2	2	2	0	$v_3$	0.500			
0	0	2	1	0	1	1	2	$v_4$	0.500			
0	1	2	1	1	1	1	1	$v_5$	0.375			
Coalition $\{ag_2, ag_7\}$												
$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_{10}$	$a_{11}$	$a_{12}$	$d$	$s(x_i, x)$			
0	1	?	?	?	2	2	2	$v_1$	0.800			
1	0	0	2	2	?	?	?	$v_1$	0.600			
1	1	?	?	?	1	1	2	$v_2$	0.400			
2	0	1	2	2	?	?	?	$v_2$	0.600			
2	0	?	?	?	0	2	2	$v_3$	0.600			

Table 10 (continued)

Coalition $\{ag_2, ag_7\}$										
2	2	2	2	?	?	?	?	?	$v_3$	0.600
2	1	0	1	1	2	2	2	0	$v_4$	0.625
2	1	1	1	1	2	2	2	1	$v_5$	0.500

**Table 11** Vectors of coalitions—Approach 2

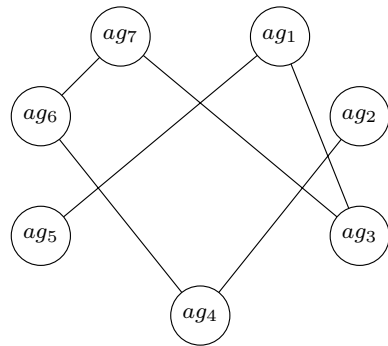
Coalition	Vector	Vector transformation	$SD_j^x$ — standard deviation
$\{ag_1, ag_3, ag_5\}$	[0.500,0.500,0.700,0.500,1]	[0.156,0.156,0.219,0.156,0.313]	0.062
$\{ag_1, ag_5, ag_7\}$	[0.800,0.500,0.600,0.750,0.900]	[0.225,0.141,0.169,0.211,0.254]	0.040
$\{ag_1, ag_6, ag_7\}$	[0.875,0.375,0.625,0.750,0.818]	[0.254,0.109,0.182,0.218,0.238]	0.052
$\{ag_2, ag_4\}$	[0.625,0.625,0.500,0.500,0.375]	[0.238,0.238,0.190,0.190,0.143]	0.036
$\{ag_2, ag_7\}$	[0.800,0.600,0.600,0.625,0.500]	[0.256,0.192,0.192,0.200,0.160]	0.031

**Table 12** Coalitions' weights and weighted average of the vectors—Approach 2

Coalition	Weight	Vector
First method—coefficients $m_{ag}^x$		
$\{ag_1, ag_3, ag_5\}$	1.833/7=0.262	[0.041,0.041,0.057,0.041,0.082]
$\{ag_1, ag_5, ag_7\}$	1.167/7=0.167	[0.038,0.023,0.028,0.035,0.042]
$\{ag_1, ag_6, ag_7\}$	1.667/7=0.238	[0.061,0.026,0.043,0.052,0.057]
$\{ag_2, ag_4\}$	1.5/7=0.214	[0.051,0.051,0.041,0.041,0.031]
$\{ag_2, ag_7\}$	0.833/7=0.119	[0.030,0.023,0.023,0.024,0.019]
Sum		[0.220,0.164,0.192,0.193,0.230]
Second method—coefficients $sd_{ag}^x$		
$\{ag_1, ag_3, ag_5\}$	0.299	[0.047,0.047,0.065,0.047,0.093]
$\{ag_1, ag_5, ag_7\}$	0.238	[0.054,0.033,0.040,0.050,0.060]
$\{ag_1, ag_6, ag_7\}$	0.238	[0.060,0.026,0.043,0.052,0.056]
$\{ag_2, ag_4\}$	0.128	[0.030,0.030,0.024,0.024,0.018]
$\{ag_2, ag_7\}$	0.098	[0.025,0.019,0.019,0.020,0.016]
Sum		[0.216,0.155,0.192,0.193,0.244]
Third method—coefficients $m_{ag}^x \cdot sd_{ag}^x$		
$\{ag_1, ag_3, ag_5\}$	0.342	[0.053,0.053,0.075,0.053,0.107]
$\{ag_1, ag_5, ag_7\}$	0.163	[0.037,0.023,0.028,0.034,0.041]
$\{ag_1, ag_6, ag_7\}$	0.235	[0.060,0.026,0.043,0.051,0.056]
$\{ag_2, ag_4\}$	0.192	[0.046,0.046,0.037,0.037,0.027]
$\{ag_2, ag_7\}$	0.068	[0.017,0.013,0.013,0.014,0.011]
Sum		[0.213,0.161,0.195,0.189,0.242]
Fourth method—coefficients $SD_j^x$		
$\{ag_1, ag_3, ag_5\}$	0.279	[0.044,0.044,0.061,0.044,0.087]
$\{ag_1, ag_5, ag_7\}$	0.183	[0.041,0.026,0.031,0.039,0.046]
$\{ag_1, ag_6, ag_7\}$	0.235	[0.060,0.026,0.043,0.051,0.056]
$\{ag_2, ag_4\}$	0.162	[0.039,0.039,0.031,0.031,0.023]
$\{ag_2, ag_7\}$	0.142	[0.036,0.027,0.027,0.028,0.023]
Sum		[0.219,0.161,0.193,0.193,0.235]

**Table 13** Function values—  
Approach 3

	$ag_1$	$ag_2$	$ag_3$	$ag_4$	$ag_5$	$ag_6$	$ag_7$
$ag_1$							
$ag_2$	1						
$ag_3$	0.4	1					
$ag_4$	1	0.4	0.8				
$ag_5$	0.4	1	0.8	0.8			
$ag_6$	0.8	0.6	0.6	0.2	0.8		
$ag_7$	0.6	0.6	0.4	0.6	0.8	0.4	

**Fig. 3** A graphical representation—Approach 3

function values are shown in Table 5. For example, the value  $\rho_A(ag_1, ag_4)$  is calculated as follows

$$\rho_A(ag_1, ag_4) = \frac{\text{card}\{v_1, v_2, v_5\}}{\text{card}\{A\}} = \frac{3}{5},$$

because values for  $v_1, v_2, v_5$  and  $ag_1, ag_4$  are different in the information system. Then the relations between the agents are determined (allied, conflict and neutrality). A graphical representation of the conflict situation is shown in Fig. 1. In this representation classifiers are circles and if classifiers are allied then circles are linked. Coalitions are the subset of vertices such that every two vertices are linked. There are three coalitions  $\{ag_1, ag_3, ag_5, ag_6, ag_7\}$ ,  $\{ag_2, ag_4\}$  and  $\{ag_4, ag_6\}$ . In the next step for each coalition an aggregated decision table is generated. For this purpose the method of approximation aggregation is used, we assume that  $m_2 = 1$ . This means that one object with the greatest similarity with the test object is selected from each decision class of the local decision table. Then objects from the same decision class that have consistent values on common attributes are written as one object in the aggregated table. The aggregated tables are presented in Table 6. For example, the third object from the aggregated table for the coalition  $\{ag_1, ag_3, ag_5, ag_6, ag_7\}$  was created by combining objects  $x_1^1 \in U_{ag_1}$ ,  $x_1^3 \in U_{ag_3}$  and  $x_1^6 \in U_{ag_6}$ , because they are from the same decision class. For attributes  $a_{12}$  and  $a_{13}$ , we have values ?, because it was not possible to select objects from the universe  $U_{ag_5}$  and  $U_{ag_7}$  and decision class

**Table 14** Aggregated decision tables—Approach 3

Coalition $\{ag_1, ag_5\}$											
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$d$	$s(x_i, x)$
1	1	1	2	1	1	0	0	2	0	$v_1$	0.500
2	2	2	2	1	2	1	2	0	2	$v_2$	0.400
1	2	0	2	2	2	2	1	2	1	$v_3$	0.700
0	0	0	1	1	0	2	1	0	2	$v_4$	0.500
1	1	0	2	1	2	2	1	2	2	$v_5$	1
Coalition $\{ag_1, ag_5\}$											
$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_{13}$	$a_{14}$	$d$	$s(x_i, x)$			
?	?	0	?	?	2	2	$v_1$	0.667			
1	1	1	2	1	?	?	$v_1$	0.800			
?	?	1	?	?	1	2	$v_2$	0.667			
2	2	2	2	1	?	?	$v_2$	0.400			
?	?	2	?	?	1	1	$v_3$	0.333			
1	2	0	2	2	?	?	$v_3$	0.600			
0	0	0	1	1	1	0	$v_4$	0.429			
1	1	0	2	1	1	2	$v_5$	1			
Coalition $\{ag_2, ag_4\}$											
$a_1$	$a_2$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_{12}$	$d$	$s(x_i, x)$		
2	1	1	0	0	2	2	2	$v_1$	0.625		
1	2	2	0	1	2	2	2	$v_2$	0.625		
1	0	2	2	2	2	2	0	$v_3$	0.500		
0	0	2	1	0	1	1	2	$v_4$	0.500		

**Table 14** (continued)

Coalition $\{ag_2, ag_4\}$									
$a_1$	$a_2$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_{12}$	$d$	$s(x_i, x)$
0	1	2	1	1	1	1	1	$v_5$	0.375
Coalition $\{ag_3, ag_7\}$									
$a_4$	$a_5$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{12}$	$d$	$s(x_i, x)$
0	1	?	?	?	2	2	2	$v_1$	0.800
?	?	1	0	0	2	0	?	$v_1$	0.200
1	1	?	?	?	1	1	2	$v_2$	0.400
?	?	2	1	2	0	2	?	$v_2$	0.400
2	0	?	?	?	0	2	2	$v_3$	0.600
2	?	?	2	1	2	1	?	$v_3$	0.800
2	1	?	?	?	2	2	0	$v_4$	0.800
0	?	?	2	1	0	2	?	$v_4$	0.600
2	1	2	2	1	2	2	1	$v_5$	0.875
Coalition $\{ag_6, ag_7\}$									
$a_4$	$a_5$	$a_6$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{14}$	$a_{15}$	$d$	$s(x_i, x)$
0	1	0	2	2	2	2	1	$v_1$	0.875
1	1	1	1	1	2	1	1	$v_2$	0.375
2	0	0	0	2	2	2	0	$v_3$	0.625
2	1	0	2	2	0	2	2	$v_4$	0.750
2	1	2	2	2	1	2	1	$v_5$	0.750

**Table 14** (continued)

Coalition $\{ag_4, ag_6\}$							
$a_1$	$a_2$	$a_6$	$a_{12}$	$a_{14}$	$a_{15}$	$d$	$s(x_i, x)$
2	1	0	2	2	1	$v_1$	0.833
1	2	1	2	1	1	$v_2$	0.500
1	0	0	0	2	0	$v_3$	0.500
0	0	0	2	2	2	$v_4$	0.500
0	1	2	1	2	1	$v_5$	0.500



**Table 15** Vectors of coalitions—Approach 3

Coalition	Vector	Vector transformation	$SD_j^x$ — standard deviation
$\{ag_1, ag_3\}$	[0.500,0.400,0.700,0.500,1]	[0.161,0.129,0.226,0.161,0.323]	0.069
$\{ag_1, ag_5\}$	[0.800,0.667,0.600,0.429,1]	[0.229,0.191,0.172,0.123,0.286]	0.055
$\{ag_2, ag_4\}$	[0.625,0.625,0.500,0.500,0.375]	[0.238,0.238,0.190,0.190,0.143]	0.036
$\{ag_3, ag_7\}$	[0.800,0.400,0.800,0.800,0.875]	[0.218,0.109,0.218,0.218,0.238]	0.046
$\{ag_6, ag_7\}$	[0.875,0.375,0.625,0.750,0.750]	[0.259,0.111,0.185,0.222,0.222]	0.050
$\{ag_4, ag_6\}$	[0.833,0.500,0.500,0.500,0.500]	[0.294,0.176,0.176,0.176,0.176]	0.047

$v_1$  that would have matching values on common attributes. Then for each aggregated table a vector of values is determined. The vector's coordinate for a given decision value is equal to the maximum similarity of objects from the decision class and the aggregated table to the test object. The values of the similarity measure are given in Table 6. When we calculate these values, we ignore the attributes for which we have value ?. The vectors  $[\mu_{j,1}(x), \dots, \mu_{j,5}(x)]$ ,  $j \in \{1, \dots, 7\}$  that were determined based on the similarity values and those which were obtained after the transformation (so that the sum of the coefficients will be equal to 1) are shown in Table 7. Then the weights for the coalitions are calculated in one of four ways and the weighted sum of the vectors is calculated.

In the first method, the coefficients  $m_{ag}^x$  are calculated for each classifier (Formula 5). We have  $m_{ag_4}^x = m_{ag_6}^x = \frac{1}{2}$ , because these classifiers belong to two coalitions. For other classifiers, the coefficient is equal to 1. Thus, the coefficients for the first, second and third coalition are respectively equal to 4.5, 1.5 and 1. The vectors of coalitions that were multiplied by the weights and the weighted average of the vectors are given in Table 8.

In the second method, the standard deviations of the vectors' coordinates that were generated by classifiers are calculated. These values are given in Table 3. The coefficients for the first, second and third coalition that were determined in accordance with Formula 6 are respectively equal to 0.595, 0.172 and 0.233. The vectors of coalitions that were multiplied by the weights and the weighted average of the vectors are given in Table 8.

In the third method, the weights for coalitions are determined according to Formula 7. This method is a combination of the two previous methods. The weights are equal to 0.704, 0.145, 0.152 for the first, second and third coalition respectively. The vectors of coalitions that were multiplied by the weights and the weighted average of the vectors are given in Table 8.

In the last method, the standard deviations of the vectors generated by the coalitions are calculated (Table 7). The coefficients for the first, second and third coalition that were determined in accordance with Formula 8 are respectively equal to 0.295, 0.401 and 0.304. The vectors of coalitions that were multiplied by the weights and the weighted average of the vectors are given in Table 8.

**Table 16** Coalitions' weights and weighted average of the vectors—Approach 3

Coalition	Weight	Vector
First method—coefficients $m_{ag}^x$		
$\{ag_1, ag_3\}$	$1/7=0.143$	[0.023,0.018,0.032,0.023,0.046]
$\{ag_1, ag_5\}$	$1.5/7=0.214$	[0.049,0.041,0.037,0.026,0.061]
$\{ag_2, ag_4\}$	$1.5/7=0.214$	[0.051,0.051,0.041,0.041,0.031]
$\{ag_3, ag_7\}$	$1/7=0.143$	[0.031,0.016,0.031,0.031,0.034]
$\{ag_6, ag_7\}$	$1/7=0.143$	[0.037,0.016,0.026,0.032,0.032]
$\{ag_4, ag_6\}$	$1/7=0.143$	[0.042,0.025,0.025,0.025,0.025]
Sum		[0.233,0.167,0.193,0.178,0.229]
Second method—coefficients $sd_{ag}^x$		
$\{ag_1, ag_3\}$	0.212	[0.034,0.027,0.048,0.034,0.069]
$\{ag_1, ag_5\}$	0.173	[0.040,0.033,0.030,0.021,0.050]
$\{ag_2, ag_4\}$	0.125	[0.030,0.030,0.024,0.024,0.018]
$\{ag_3, ag_7\}$	0.180	[0.039,0.020,0.039,0.039,0.043]
$\{ag_6, ag_7\}$	0.140	[0.036,0.016,0.026,0.031,0.031]
$\{ag_4, ag_6\}$	0.169	[0.050,0.030,0.030,0.030,0.030]
Sum		[0.229,0.155,0.197,0.180,0.240]
Third method—coefficients $m_{ag}^x \cdot sd_{ag}^x$		
$\{ag_1, ag_3\}$	0.190	[0.031,0.025,0.043,0.031,0.061]
$\{ag_1, ag_5\}$	0.227	[0.052,0.043,0.039,0.028,0.065]
$\{ag_2, ag_4\}$	0.145	[0.034,0.034,0.028,0.028,0.021]
$\{ag_3, ag_7\}$	0.161	[0.035,0.018,0.035,0.035,0.038]
$\{ag_6, ag_7\}$	0.126	[0.033,0.014,0.023,0.028,0.028]
$\{ag_4, ag_6\}$	0.152	[0.045,0.027,0.027,0.027,0.027]
Sum		[0.229,0.161,0.195,0.176,0.240]
Fourth method—coefficients $SD_j^x$		
$\{ag_1, ag_3\}$	0.227	[0.037,0.029,0.051,0.037,0.073]
$\{ag_1, ag_5\}$	0.181	[0.042,0.035,0.031,0.022,0.052]
$\{ag_2, ag_4\}$	0.118	[0.028,0.028,0.022,0.022,0.017]
$\{ag_3, ag_7\}$	0.153	[0.033,0.017,0.033,0.033,0.036]
$\{ag_6, ag_7\}$	0.166	[0.043,0.018,0.031,0.037,0.037]
$\{ag_4, ag_6\}$	0.155	[0.046,0.027,0.027,0.027,0.027]
Sum		[0.228,0.154,0.196,0.179,0.243]

It should be noted that each method of calculating weights generates completely different values. In the first, second and third methods, the first coalition  $\{ag_1, ag_3, ag_5, ag_6, ag_7\}$  was recognized as the strongest. However, there are differences in the ranking of remaining coalitions. The last method of calculating weights recognized the second coalition  $\{ag_2, ag_4\}$  as the strongest.

After calculating the weighted sum of vectors according to one of the methods discussed above (see the last line in Table 8), the decision with the highest

**Table 17** Example—comparison of results

Approach	Coalitions	Method of calculating weights	Result	
			$\varepsilon = 0.019$	$\varepsilon = 0.01$
Approach 1	$\{ag_1, ag_3, ag_5, ag_6, ag_7\}$ , $\{ag_2, ag_4\}$ , $\{ag_4, ag_6\}$	1	$\{v_1\}$	$\{v_1\}$
		2	$\{v_1\}$	$\{v_1\}$
		3	$\{v_1, v_3, v_4, v_5\}$	$\{v_1\}$
		4	$\{v_1\}$	$\{v_1\}$
Approach 2	$\{ag_1, ag_3, ag_5\}$ , $\{ag_1, ag_5, ag_7\}$ , $\{ag_1, ag_6, ag_7\}$ , $\{ag_2, ag_4\}, \{ag_2, ag_7\}$	1	$\{v_1, v_5\}$	$\{v_1, v_5\}$
		2	$\{v_5\}$	$\{v_5\}$
		3	$\{v_5\}$	$\{v_5\}$
		4	$\{v_1, v_5\}$	$\{v_5\}$
Approach 3	$\{ag_1, ag_3\}, \{ag_1, ag_5\}$ , $\{ag_6, ag_7\}$ , $\{ag_4, ag_6\}$	1	$\{v_1, v_5\}$	$\{v_1, v_5\}$
		2	$\{v_1, v_5\}$	$\{v_5\}$
		3	$\{v_1, v_5\}$	$\{v_5\}$
		4	$\{v_1, v_5\}$	$\{v_5\}$

coefficient is determined and its group is generated using the DBSCAN algorithm. In each method, the decision  $v_1$  has the highest vector's coefficient. However, comparing the vectors generated in each of the methods we can see the differences that affect the result in the DBSCAN method. For example, for the parameters value  $\varepsilon = 0.019$  and  $minPts = 2$  of the DBSCAN algorithm, the third method generates the set  $\{v_1, v_3, v_4, v_5\}$  while all others methods generate the set  $\{v_1\}$ . For the parameters value  $\varepsilon = 0.01$  and  $minPts = 2$  of the DBSCAN algorithm, all the methods determine the set  $\{v_1\}$ .

## 5.2 Approach 2

In *Approach 2* we use the second method of defining an information system (Formula 2) and the function of distance between agents (Formula 3). The function values are shown in Table 9. For example, the value  $\rho_A^*(ag_1, ag_4)$  is calculated as follows

$$\rho_A^*(ag_1, ag_4) = \frac{0.5 + 1 + 0.5 + 0.5 + 0.5}{card\{A\}} = \frac{3}{5},$$

because  $\phi_{v_1}^*(ag_1, ag_4) = \phi_{v_3}^*(ag_1, ag_4) = \phi_{v_4}^*(ag_1, ag_4) = \phi_{v_5}^*(ag_1, ag_4) = 0.5$  and  $\phi_{v_2}^*(ag_1, ag_4) = 1$ . Similarly as before, a graphical representation of the conflict situation was prepared (Fig. 2). As can be seen, there are five coalitions  $\{ag_1, ag_3, ag_5\}$ ,  $\{ag_1, ag_5, ag_7\}$ ,  $\{ag_1, ag_6, ag_7\}$ ,  $\{ag_2, ag_4\}$  and  $\{ag_2, ag_7\}$ .

For each coalition an aggregated decision table is generated (see Table 10). The vectors that were determined for each aggregated table are given in Table 11. Then the weights for coalitions are calculated in one of four ways and the weighted sums of the vectors are calculated. These values are given in Table 12.

This time the decision  $v_5$  received the greatest support of the coalitions, differently than for Approach 1. Of course, also completely different vectors were

**Table 18** Comparison the methods of determining the strength of the coalition (Approach 1 and Soybean)

System	Weights	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	1/3/0.0031	<b>0.021</b>	0.271	1.886				
	2	1/3/0.0031	<b>0.021</b>	0.271	1.886				
	3	1/3/0.0031	<b>0.021</b>	0.271	1.886				
	4	1/3/0.0031	<b>0.021</b>	0.271	1.886				
$DS_{Ag5}^S$	1	8/2/0.0028	<b>0.032</b>	0.295	2.093	8/2/0.0021	<b>0.053</b>	0.255	1.529
	2	8/2/0.0026	0.035	0.295	2.037	8/2/0.0021	<b>0.053</b>	0.255	1.516
	3	8/2/0.0027	0.035	0.295	2.027	8/2/0.0021	<b>0.053</b>	0.255	1.524
	4	8/2/0.0023	0.035	0.290	2.043	8/2/0.0016	0.059	0.234	1.476
$DS_{Ag7}^S$	1	10/5/0.003	0.053	0.383	1.965	10/5/0.0018	<b>0.069</b>	0.316	1.529
	2	10/5/0.0026	<b>0.051</b>	0.375	2.112	10/5/0.0016	0.074	0.309	1.521
	3	10/5/0.0029	0.053	0.386	2.085	10/5/0.0017	0.072	0.314	1.529
	4	10/5/0.0017	0.069	0.338	2.005	10/5/0.0011	0.077	0.311	1.564
$DS_{Ag9}^S$	1	6/2/0.0027	0.043	0.351	2.000	6/2/0.0017	0.061	0.301	1.548
	2	6/2/0.0026	0.045	0.343	1.963	6/2/0.0015	0.064	0.285	1.492
	3	6/2/0.0028	0.037	0.354	2.112	6/2/0.0016	0.064	0.285	1.495
	4	1/3/0.0024	<b>0.032</b>	0.391	1.984	1/3/0.0015	<b>0.045</b>	0.319	1.521
$DS_{Ag11}^S$	1	7/2/0.0024	0.109	0.426	1.968	7/2/0.0016	0.144	0.388	1.540
	2	7/2/0.0024	<b>0.104</b>	0.420	1.880	7/2/0.0015	<b>0.138</b>	0.383	1.559
	3	7/2/0.0024	<b>0.104</b>	0.420	1.894	7/2/0.0016	<b>0.138</b>	0.386	1.551
	4	4/3/0.0015	0.136	0.410	1.904	4/3/0.0009	0.144	0.391	1.561

obtained for all four methods of determining the coalitions' weights. As a consequence, we get different results. For the parameters value  $\epsilon = 0.019$  and  $minPts = 2$  of the DBSCAN algorithm, the first and the last methods generate the set  $\{v_1, v_5\}$  while the second and the third methods generate the set  $\{v_5\}$ . For the parameters value  $\epsilon = 0.01$  and  $minPts = 2$  of the DBSCAN algorithm, only the first method determines the set  $\{v_1, v_5\}$ , while the rest of the methods generate the set  $\{v_5\}$ .

### 5.3 Approach 3

In *Approach 3* we use the second method of defining an information system (Formula 2) and the conflict function (Formula 4). The function values are shown in Table 13 and a graphical representation of the conflict situation is presented in Fig. 3. As can be seen, there are six coalitions  $\{ag_1, ag_3\}$ ,  $\{ag_1, ag_5\}$ ,  $\{ag_2, ag_4\}$ ,  $\{ag_3, ag_7\}$ ,  $\{ag_6, ag_7\}$  and  $\{ag_4, ag_6\}$ .

The aggregated decision tables for the coalitions are given in Table 14 and the vectors that were generated based on the aggregated tables are presented in Table 15. The coalitions' weights and the weighted sum of the vectors are given in Table 16.

The results obtained for Approach 3 are significantly different from the two previously considered approaches. This time the decision  $v_1$  received the greatest support

**Table 19** Comparison the methods of determining the strength of the coalition (Approach 2 and Soybean)

System	Weights	$m_1/m_2/\varepsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\varepsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	2/3/0.0031	<b>0.021</b>	0.271	1.886				
	2	2/3/0.0031	<b>0.021</b>	0.271	1.886				
	3	2/3/0.0031	<b>0.021</b>	0.271	1.886				
	4	2/3/0.0031	<b>0.021</b>	0.271	1.886				
$DS_{Ag5}^S$	1	9/2/0.00273	0.037	0.290	2.013	9/2/0.0021	<b>0.053</b>	0.255	1.508
	2	9/2/0.00259	<b>0.035</b>	0.295	2.043	9/2/0.0021	<b>0.053</b>	0.253	1.468
	3	9/2/0.00266	0.037	0.290	2.024	9/2/0.0021	<b>0.053</b>	0.253	1.473
	4	9/2/0.00238	<b>0.035</b>	0.290	2.040	9/2/0.00168	0.056	0.245	1.535
$DS_{Ag7}^S$	1	4/6/0.0029	<b>0.040</b>	0.356	2.080	4/6/0.00196	0.064	0.314	1.513
	2	10/10/0.00266	0.045	0.346	2.093	10/10/0.00196	<b>0.053</b>	0.311	1.556
	3	4/6/0.0028	0.043	0.348	2.069	4/6/0.00175	0.064	0.311	1.527
	4	6/6/0.00224	0.051	0.367	1.989	6/6/0.00147	0.064	0.314	1.543
$DS_{Ag9}^S$	1	8/2/0.00238	0.035	0.324	1.902	8/2/0.00168	0.051	0.279	1.545
	2	8/2/0.00266	0.032	0.327	2.069	8/2/0.00154	<b>0.043</b>	0.276	1.556
	3	8/2/0.00273	<b>0.029</b>	0.335	2.082	8/2/0.00161	0.048	0.274	1.532
	4	10/2/0.00168	0.032	0.319	1.955	10/2/0.00091	0.051	0.274	1.492
$DS_{Ag11}^S$	1	4/2/0.00245	0.066	0.370	1.785	4/2/0.00168	0.085	0.330	1.551
	2	4/1/0.0028	<b>0.056</b>	0.420	2.104	4/1/0.00168	<b>0.077</b>	0.338	1.527
	3	4/1/0.00259	0.059	0.386	1.856	4/1/0.00168	0.080	0.332	1.511
	4	4/3/0.00182	0.072	0.394	1.785	4/3/0.00119	<b>0.077</b>	0.338	1.537

for the first method of calculating weights, for rest of the methods the decision  $v_5$  received the biggest support. We get different results than before. For the parameters value  $\varepsilon = 0.019$  and  $minPts = 2$  of the DBSCAN algorithm, all the methods of calculating weights generate the set  $\{v_1, v_5\}$ . For the parameters value  $\varepsilon = 0.01$  and  $minPts = 2$  of the DBSCAN algorithm, the first method determines the set  $\{v_1, v_5\}$ , while rest of the methods generate the set  $\{v_5\}$ .

A comparison of the results obtained using three approaches to creating coalitions and four methods of calculating coalitions weights is presented in Table 17. As can be seen, *Approach 1* determined the largest coalitions, while *Approach 3* generated the smallest coalitions. The decisions  $v_1$  and  $v_5$  are preferred by all methods. However, the exact indication depends on the approach and the method of calculating the weights. Even within one approach, the use of weights can affect that another decision has the greatest support of classifiers. It can be generalized that *Approach 1* considers that the best decision is the decision  $v_1$ , while *Approaches 2 and 3* recognize that the best decision is the decision  $v_5$ .

**Table 20** Comparison the methods of determining the strength of the coalition (Approach 3 and Soybean)

System	Weights	$m_1/m_2/\varepsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\varepsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	2/3/0.0031	<b>0.021</b>	0.271	1.886				
	2	2/3/0.0031	<b>0.021</b>	0.271	1.886				
	3	2/3/0.0031	<b>0.021</b>	0.271	1.886				
	4	2/3/0.0031	<b>0.021</b>	0.271	1.886				
$DS_{Ag5}^S$	1	1/2/0.0027	<b>0.016</b>	0.303	2.106	1/2/0.0017	<b>0.032</b>	0.242	1.439
	2	1/2/0.0027	0.021	0.306	2.088	1/2/0.0016	0.040	0.245	1.439
	3	1/2/0.0027	0.019	0.311	2.011	1/2/0.0014	0.040	0.234	1.386
	4	1/2/0.00259	0.024	0.324	2.077	1/2/0.0014	0.051	0.239	1.410
$DS_{Ag7}^S$	1	1/2/0.0022	<b>0.013</b>	0.322	1.793	1/2/0.0017	<b>0.027</b>	0.293	1.574
	2	1/2/0.0021	0.016	0.338	1.832	1/2/0.0014	0.029	0.295	1.519
	3	1/2/0.0019	<b>0.013</b>	0.324	1.771	1/2/0.0014	0.029	0.285	1.516
	4	5/6/0.00147	0.037	0.324	1.936	5/6/0.00105	0.048	0.285	1.564
$DS_{Ag9}^S$	1	2/3/0.0027	0.021	0.338	2.066	2/3/0.0017	0.053	0.247	1.319
	2	2/3/0.0027	0.035	0.332	2.029	2/3/0.0016	0.061	0.245	1.311
	3	2/3/0.0026	0.035	0.332	2.008	2/3/0.0017	0.053	0.253	1.351
	4	1/3/0.00245	<b>0.011</b>	0.335	1.973	1/3/0.00119	<b>0.045</b>	0.255	1.346
$DS_{Ag11}^S$	1	2/4/0.003	0.029	0.311	1.883	2/4/0.0022	0.037	0.277	1.511
	2	2/2/0.0027	0.029	0.332	1.777	2/2/0.0022	0.037	0.293	1.505
	3	2/2/0.0024	<b>0.027</b>	0.314	1.593	2/2/0.0021	<b>0.035</b>	0.279	1.505
	4	2/5/0.00203	0.040	0.407	1.894	2/5/0.00133	0.064	0.314	1.471

## 6 Experimental Protocol

In this section, a description of the experiments that were carried out is included. Datasets and evaluation measures are presented.

Three datasets from the UCI repository were selected for the experiments: Soybean, Vehicle Silhouettes and Landsat Satellite. These sets were randomly dispersed into five different versions. The version with three, five, seven, nine and eleven local decision tables was considered. The dispersion was not intended to increase system efficiency but to provide dispersed knowledge, that is why the process was not optimized in any way. These sets were also used in the previous work (Przybyła-Kasperek and Wakulicz-Deja 2014a, b, 2016), where you can find a detailed description of the data used and the process of dispersion.

We use the following designations for dispersed systems:

- $DS_{Ag3}^S$ —dispersed system with 3 tables,
- $DS_{Ag5}^S$ —dispersed system with 5 tables,
- $DS_{Ag7}^S$ —dispersed system with 7 tables,
- $DS_{Ag9}^S$ —dispersed system with 9 tables,
- $DS_{Ag11}^S$ —dispersed system with 11 tables.

**Table 21** Comparison of the efficiency of inference using Approaches 1, 2, 3 and without the strength of the coalition (Soybean)

System	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{a}_{DS}$	$t$
Approach 1					
$DS_{Ag3}$	1/6/0.0043	0.021	0.295	1.939	0.17
$DS_{Ag5}$	6/2/0.0043	0.051	0.290	1.973	0.10
	6/2/0.0024	0.061	0.253	1.484	0.10
$DS_{Ag7}$	10/6/0.005	0.059	0.378	2.027	2.58
	10/6/0.0032	0.069	0.327	1.561	2.58
$DS_{Ag9}$	6/3/0.0044	0.048	0.340	1.997	2.27
	6/3/0.0024	0.064	0.309	1.559	2.27
$DS_{Ag11}$	8/3/0.00425	0.106	0.399	2.037	5.25
	8/3/0.0023	0.141	0.378	1.481	5.25
Approach 2					
$DS_{Ag3}$	1/6/0.0043	<b>0.021</b>	0.295	1.861	0.16
$DS_{Ag5}$	9/2/0.003	0.053	0.261	1.564	0.10
$DS_{Ag7}$	4/10/0.0044	0.059	0.319	1.676	5.36
	4/10/0.0041	0.061	0.311	1.535	5.36
$DS_{Ag9}$	5/3/0.0026	<b>0.037</b>	0.316	1.718	2.19
	5/3/0.0023	<b>0.048</b>	0.279	1.527	2.19
$DS_{Ag11}$	4/2/0.00575	0.056	0.407	1.992	1.55
	4/2/0.00425	0.074	0.319	1.516	1.55
Approach 3					
$DS_{Ag3}$	1/3/0.0041	0.021	0.277	1.883	0.08
$DS_{Ag5}$	1/2/0.00605	<b>0.019</b>	0.290	2.045	0.07
	1/2/0.00365	<b>0.048</b>	0.245	1.481	0.07
$DS_{Ag7}$	1/10/0.0082	<b>0.019</b>	0.362	2.003	0.46
	1/10/0.006	<b>0.029</b>	0.306	1.585	0.46
$DS_{Ag9}$	6/3/0.0066	0.045	0.364	2.011	2.05
	6/3/0.004	0.066	0.279	1.383	2.05
$DS_{Ag11}$	2/5/0.0101	<b>0.045</b>	0.348	1.819	7.11
	2/5/0.0077	<b>0.056</b>	0.293	1.505	7.11

In this paper, the quality of classification of a dispersed system with three different approaches to conflict analysis and four different methods of determining the strength of a coalition is examined. For this purpose, the train and test methodology was used, since it is suitable for dispersed data, as was explained in detail in Przybyła-Kasperek and Wakulicz-Deja (2016). For the Landsat Satellite data set validation, training, and test sets were used.

Because the method of density-based algorithm that was used generates a set of decisions rather than a single decision, so the appropriate measures of quality of classification are used. There are:

- estimator of classification error

**Table 22** Comparison the methods of determining the strength of the coalition (Approach 1 and Vehicle Silhouettes)

System	Weights	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	1/9/0.0021	<b>0.118</b>	0.476	1.520	1/9/0.0009	<b>0.185</b>	0.366	1.217
	2	1/9/0.0022	0.122	0.508	1.543	1/9/0.001	0.201	0.374	1.213
	3	1/9/0.0022	0.122	0.508	1.543	1/9/0.001	0.201	0.374	1.213
	4	2/9/0.0022	0.134	0.531	1.539	2/9/0.001	0.213	0.417	1.236
$DS_{Ag5}^S$	1	3/8/0.0016	0.169	0.524	1.516	3/8/0.0009	<b>0.205</b>	0.413	1.256
	2	3/8/0.0018	<b>0.157</b>	0.539	1.535	3/8/0.0009	0.217	0.406	1.228
	3	3/8/0.0018	<b>0.157</b>	0.539	1.535	3/8/0.0009	0.217	0.406	1.228
	4	3/10/0.0016	0.240	0.610	1.524	3/10/0.0007	0.327	0.508	1.244
$DS_{Ag7}^S$	1	4/8/0.0019	<b>0.094</b>	0.465	1.531	4/8/0.0007	0.197	0.378	1.232
	2	5/8/0.0019	0.098	0.480	1.551	5/8/0.0008	<b>0.177</b>	0.370	1.240
	3	5/8/0.0019	0.098	0.480	1.551	5/8/0.0008	<b>0.177</b>	0.370	1.240
	4	1/5/0.0014	0.189	0.594	1.555	1/5/0.0007	0.272	0.480	1.256
$DS_{Ag9}^S$	1	5/4/0.001	<b>0.126</b>	0.528	1.555	5/4/0.0005	<b>0.224</b>	0.421	1.248
	2	5/4/0.0011	0.142	0.512	1.528	5/4/0.0005	0.240	0.425	1.224
	3	5/5/0.0012	0.130	0.520	1.551	5/5/0.0005	0.244	0.429	1.232
	4	4/5/0.001	0.193	0.591	1.579	4/5/0.0005	0.276	0.461	1.280
$DS_{Ag11}^S$	1	2/2/0.0009	<b>0.122</b>	0.480	1.524	2/2/0.0004	<b>0.185</b>	0.354	1.224
	2	3/1/0.0009	0.134	0.512	1.512	3/1/0.0005	0.193	0.398	1.244
	3	3/1/0.001	0.146	0.528	1.539	3/1/0.0005	0.193	0.382	1.228
	4	2/3/0.0008	0.181	0.567	1.535	2/3/0.0003	0.248	0.437	1.217

$$e = \frac{1}{card\{U_{test}\}} \sum_{x \in U_{test}} I(d(x) \notin \hat{d}_{DS}(x)),$$

where  $I(d(x) \notin \hat{d}_{DS}(x)) = 1$ , when  $d(x_i) \notin \hat{d}_{DS}(x)$  and  $I(d(x) \notin \hat{d}_{DS}(x)) = 0$ , when  $d(x) \in \hat{d}_{DS}(x)$ ;  $\hat{d}_{DS}(x)$  is a set of global decisions generated by the system for the test object  $x$

- estimator of classification ambiguity error

$$e_{ONE} = \frac{1}{card\{U_{test}\}} \sum_{x \in U_{test}} I(d(x) \neq \hat{d}_{DS}(x)),$$

where  $I(d(x) \neq \hat{d}_{DS}(x)) = 1$ , when  $\{d(x)\} \neq \hat{d}_{DS}(x)$  and  $I(d(x) \neq \hat{d}_{DS}(x)) = 0$ , when  $\{d(x)\} = \hat{d}_{DS}(x)$

- the average size of the global decisions sets

$$\bar{d}_{DS} = \frac{1}{card\{U_{test}\}} \sum_{x \in U_{test}} card\{\hat{d}_{DS}(x)\}.$$

In the first stage of experiments, optimization of system's parameters  $m_1$ ,  $m_2$  and  $\epsilon$  was performed. As was mentioned earlier, the parameter  $m_1$  is used in the process



**Table 23** Comparison the methods of determining the strength of the coalition (Approach 2 and Vehicle Silhouettes)

System	Weights	$m_1/m_2/\varepsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\varepsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	1/9/0.0022	<b>0.110</b>	0.469	1.539	1/9/0.001	<b>0.189</b>	0.370	1.232
	2	3/2/0.0024	<b>0.110</b>	0.484	1.520	3/2/0.0011	0.197	0.382	1.217
	3	3/2/0.0024	<b>0.110</b>	0.488	1.524	3/2/0.0011	0.197	0.382	1.217
	4	3/2/0.0022	0.130	0.496	1.520	3/2/0.0011	0.213	0.398	1.224
$DS_{Ag5}^S$	1	2/1/0.0016	<b>0.154</b>	0.543	1.555	2/1/0.0009	<b>0.224</b>	0.433	1.260
	2	7/1/0.0019	0.157	0.547	1.528	7/1/0.0009	0.236	0.449	1.260
	3	6/1/0.0017	0.167	0.547	1.520	6/1/0.0008	0.232	0.429	1.236
	4	6/1/0.0016	0.236	0.591	1.535	6/1/0.0008	0.291	0.476	1.260
$DS_{Ag7}^S$	1	2/8/0.0021	0.110	0.472	1.524	2/8/0.0008	0.177	0.362	1.232
	2	2/5/0.0022	<b>0.091</b>	0.480	1.535	2/5/0.0009	<b>0.173</b>	0.354	1.228
	3	6/6/0.0019	0.114	0.472	1.508	6/6/0.0008	0.177	0.346	1.220
	4	6/6/0.0018	0.146	0.516	1.528	6/6/0.0009	0.248	0.409	1.228
$DS_{Ag9}^S$	1	5/4/0.001	<b>0.122</b>	0.512	1.543	5/4/0.0005	<b>0.213</b>	0.378	1.220
	2	2/4/0.0012	0.126	0.524	1.535	2/4/0.0006	0.228	0.421	1.248
	3	5/4/0.0011	0.134	0.524	1.539	5/4/0.0005	0.217	0.413	1.256
	4	5/4/0.0011	0.165	0.535	1.539	5/4/0.0005	0.264	0.417	1.228
$DS_{Ag11}^S$	1	1/1/0.0009	0.122	0.484	1.520	1/1/0.0005	<b>0.177</b>	0.378	1.256
	2	1/1/0.0011	<b>0.118</b>	0.504	1.528	1/1/0.0005	0.197	0.374	1.236
	3	1/1/0.0011	0.130	0.496	1.531	1/1/0.0006	0.185	0.378	1.252
	4	1/1/0.001	0.154	0.520	1.539	1/1/0.0004	0.244	0.402	1.236

of selecting the relevant objects based on which the vectors  $\bar{\mu}_{ag}(x)$  are defined. The parameter  $m_2$  is used, analogously, when defining the aggregated tables. In order to determine the optimal values for these parameters, experiments with different parameter's values ( $m_1, m_2 \in \{1, \dots, 10\}$  for Soybean and Vehicle data sets and  $m_1, m_2 \in \{1, \dots, 5\}$  for Satellite) and the voting method (instead of the density algorithm) were performed (so that the  $\varepsilon$  parameter was not needed at this stage). Then the arguments for which the function

$$f(m_1, m_2) = 0.6 \cdot e + 0.4 \cdot \bar{d}_{DS}$$

reaches the minimum values were chosen. In this way optimization of the parameter values  $m_1$  and  $m_2$  was carried out taking into account two measures of quality of classification  $e$  and  $\bar{d}_{DS}$ . Then the value of  $\varepsilon$  was optimized. For this purpose, a number of experiments with different values of this parameter were performed. The smallest analyzed value was 0.0001, and was increased by 0.00005. Finally, the values of  $\varepsilon$  for which the greatest improvement in the quality was obtained and that ensures the average size of the global decisions sets comparable with the results presented in Przybyła-Kasperek (2017) were selected. Optimal parameter values are given in the tables with results, which are presented in the next section.

**Table 24** Comparison the methods of determining the strength of the coalition (Approach 3 and Vehicle Silhouettes)

System	Weights	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	1/7/0.0022	0.114	0.457	1.504	1/7/0.001	0.193	0.378	1.228
	2	5/10/0.0026	<b>0.110</b>	0.496	1.531	5/10/0.0012	<b>0.189</b>	0.366	1.217
	3	5/10/0.0026	<b>0.110</b>	0.496	1.531	5/10/0.0012	<b>0.189</b>	0.366	1.217
	4	2/4/0.0026	<b>0.110</b>	0.496	1.543	2/4/0.0011	0.209	0.378	1.213
$DS_{Ag5}^S$	1	1/8/0.00168	<b>0.150</b>	0.531	1.559	1/8/0.00102	<b>0.224</b>	0.433	1.264
	2	9/4/0.002	0.189	0.555	1.500	9/4/0.0011	0.248	0.433	1.236
	3	9/4/0.002	0.189	0.555	1.500	9/4/0.0011	0.248	0.433	1.236
	4	1/1/0.0018	0.213	0.571	1.504	1/1/0.0008	0.287	0.484	1.232
$DS_{Ag7}^S$	1	1/7/0.0019	0.102	0.472	1.528	1/7/0.0009	0.189	0.374	1.236
	2	5/7/0.00202	<b>0.087</b>	0.469	1.535	5/7/0.0008	<b>0.181</b>	0.335	1.217
	3	5/7/0.00202	<b>0.087</b>	0.469	1.535	5/7/0.0008	<b>0.181</b>	0.335	1.217
	4	1/8/0.0018	0.161	0.551	1.539	1/8/0.0007	0.236	0.417	1.228
$DS_{Ag9}^S$	1	5/4/0.00105	<b>0.126</b>	0.528	1.543	5/4/0.0005	<b>0.232</b>	0.421	1.248
	2	5/1/0.0011	0.146	0.516	1.535	5/1/0.0005	0.256	0.429	1.228
	3	4/2/0.0012	0.154	0.539	1.547	4/2/0.0006	0.248	0.425	1.232
	4	4/4/0.001	0.201	0.551	1.492	4/4/0.0005	0.299	0.465	1.224
$DS_{Ag11}^S$	1	1/1/0.001	<b>0.114</b>	0.500	1.535	1/1/0.0005	0.181	0.370	1.260
	2	5/3/0.001	0.130	0.480	1.496	5/3/0.0005	<b>0.177</b>	0.346	1.224
	3	5/3/0.0011	0.118	0.496	1.547	5/3/0.0005	0.181	0.350	1.224
	4	2/3/0.001	0.173	0.539	1.535	2/3/0.0005	0.217	0.394	1.236

## 7 Experiments and Discussion

In this section, the results of experiments with the division on the analyzed datasets will be presented. Then comparison of the results with regards to three conflict analysis methods and four methods of determining the strength of the coalition will be made in the last part of this section.

### 7.1 Results with the Soybean Data Set

The results of the experiments with the Soybean data set are presented in Tables 18, 19, 20. Each table illustrates approach that was used to generate the coalitions: *Approach 1, 2 and 3*. The following information is given in the tables:

- the version of dispersion (System)
- the methods of determining the strength of the coalition (Weights): 1—the size of the coalition, 2—the unambiguous of the decisions made by the classifiers,

**Table 25** Comparison of the efficiency of inference using Approaches 1, 2, 3 and without the strength of the coalition (Vehicle Silhouettes)

System	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{a}_{DS}$	$t$
Approach 1					
$DS_{Ag3}$	1/2/0.0036	0.134	0.500	1.535	0.05
	1/2/0.0018	0.197	0.425	1.272	0.05
$DS_{Ag5}$	1/2/0.0033	0.228	0.626	1.535	0.06
	1/2/0.00195	0.295	0.480	1.224	0.06
$DS_{Ag7}$	1/7/0.00405	0.201	0.614	1.547	0.47
	1/7/0.0018	0.268	0.492	1.260	0.47
$DS_{Ag9}$	10/2/0.00225	0.232	0.614	1.508	0.15
	10/2/0.0015	0.303	0.528	1.276	0.15
$DS_{Ag11}$	10/4/0.0027	0.197	0.602	1.532	5.10
	10/4/0.00135	0.252	0.449	1.236	5.10
Approach 2					
$DS_{Ag3}$	9/2/0.003	0.118	0.500	1.539	0.07
	9/2/0.0018	0.181	0.417	1.268	0.07
$DS_{Ag5}$	2/10/0.00315	0.217	0.602	1.531	0.52
	2/10/0.00135	0.283	0.465	1.220	0.52
$DS_{Ag7}$	10/6/0.0039	<b>0.146</b>	0.535	1.535	1.02
	10/6/0.0021	0.205	0.425	1.276	1.02
$DS_{Ag9}$	10/5/0.003	0.205	0.567	1.528	9.43
	10/5/0.0015	<b>0.264</b>	0.457	1.252	9.43
$DS_{Ag11}$	9/3/0.00315	0.154	0.531	1.531	2.46
	9/3/0.00135	<b>0.205</b>	0.421	1.252	2.46
Approach 3					
$DS_{Ag3}$	1/2/0.00495	<b>0.102</b>	0.472	1.535	0.04
	1/2/0.0024	<b>0.177</b>	0.390	1.256	0.04
$DS_{Ag5}$	1/9/0.0048	<b>0.177</b>	0.520	1.524	0.11
	1/9/0.0021	<b>0.264</b>	0.449	1.232	0.11
$DS_{Ag7}$	10/6/0.0063	0.154	0.516	1.500	0.21
	10/6/0.0033	<b>0.205</b>	0.413	1.244	0.21
$DS_{Ag9}$	1/1/0.0054	<b>0.189</b>	0.543	1.512	0.08
	1/1/0.00255	0.268	0.457	1.260	0.08
$DS_{Ag11}$	3/1/0.00465	<b>0.154</b>	0.520	1.520	0.40
	3/1/0.0024	0.213	0.409	1.272	0.40

3—the size of the coalition and the unambiguous of the taken decisions, 4—the unambiguous of the decisions made by the coalition

- the optimal parameters values  $m_1$ ,  $m_2$  and  $\epsilon$
- the measures to determine the quality of inference:  $e$ ,  $e_{ONE}$  and  $\bar{a}_{DS}$

In Tables 18, 19, 20 the best results were in bold.

**Table 26** Comparison the methods of determining the strength of the coalition (Approach 1 and Landsat Satellites)

System	Weights	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	1/4/0.0013	<b>0.044</b>	0.409	1.745				
	2	1/4/0.0013	<b>0.044</b>	0.409	1.745				
	3	1/4/0.0013	<b>0.044</b>	0.409	1.745				
	4	1/4/0.0013	<b>0.044</b>	0.409	1.745				
$DS_{Ag5}^S$	1	1/4/0.0015	<b>0.015</b>	0.381	1.656	1/4/0.0007	<b>0.045</b>	0.222	1.240
	2	1/4/0.0015	<b>0.015</b>	0.381	1.656	1/4/0.0007	<b>0.045</b>	0.222	1.240
	3	1/4/0.0015	<b>0.015</b>	0.381	1.656	1/4/0.0007	<b>0.045</b>	0.222	1.240
	4	1/4/0.0015	<b>0.015</b>	0.381	1.656	1/4/0.0007	<b>0.045</b>	0.223	1.241
$DS_{Ag7}^S$	1	2/4/0.0013	<b>0.015</b>	0.371	1.612	2/4/0.0006	0.054	0.226	1.237
	2	2/4/0.00135	0.016	0.384	1.647	2/4/0.0006	0.054	0.226	1.237
	3	2/4/0.00135	0.016	0.384	1.647	2/4/0.0006	0.054	0.226	1.237
	4	2/4/0.00135	<b>0.015</b>	0.384	1.650	2/4/0.0006	<b>0.053</b>	0.226	1.238
$DS_{Ag9}^S$	1	1/2/0.00105	0.035	0.404	1.612	1/2/0.0005	<b>0.078</b>	0.264	1.236
	2	1/2/0.00105	<b>0.034</b>	0.404	1.613	1/2/0.0005	0.082	0.265	1.234
	3	1/2/0.00105	0.035	0.405	1.612	1/2/0.0005	0.079	0.263	1.237
	4	1/2/0.00105	0.037	0.408	1.610	1/2/0.0005	0.086	0.270	1.233
$DS_{Ag11}^S$	1	1/1/0.00105	<b>0.022</b>	0.386	1.616	1/1/0.0005	0.063	0.231	1.223
	2	1/1/0.00105	0.024	0.383	1.607	1/1/0.0005	0.064	0.231	1.220
	3	1/1/0.00105	<b>0.022</b>	0.384	1.611	1/1/0.0005	<b>0.062</b>	0.233	1.226
	4	1/1/0.0011	0.025	0.398	1.646	1/1/0.0005	0.064	0.236	1.229

Comparing the results presented in Tables 18, 19 and 20 the following conclusions can be drawn. Approach 3 of creating coalitions of classifiers generates the best results. The improvement of the quality of classification is significant compared to Approaches 1 and 2. In general it can be said that the worst results were obtained using Approach 1. Approach 2 improves the quality of classification compared to Approach 1. However, both are much worse than Approach 3. These conclusion apply to ensemble of more than three classifiers. For three classifiers, method of conflict analysis and method of determining the strength of the coalition are irrelevant. It is easy to justify this result. Simply, three classifiers are too small to generate any meaningful coalitions. It can not be concluded that any of the methods of determining the strength of the coalition is better than others.

The results obtained using the coalitions' strength were compared with the results obtained when the weights were not used. These results are shown in Table 21. The best results within this table were in bold. It was noted that the use of coalitions' weights significantly improved the quality of classification. But the most spectacular improvement was obtained for Approach 3. Such results are not surprising as already in the paper (Przybyła-Kasperek 2017) it was shown that Approach 3 recognizes relations between classifiers in the best way and reflects them in the form of coalitions. This is why, in this approach, the strengthening of the importance of

**Table 27** Comparison the methods of determining the strength of the coalition (Approach 2 and Landsat Satellites)

System	Weights	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	2/4/0.0013	<b>0.041</b>	0.405	1.738				
	2	2/4/0.0013	<b>0.041</b>	0.404	1.734				
	3	2/4/0.0013	<b>0.041</b>	0.404	1.734				
	4	4/4/0.0013	<b>0.041</b>	0.404	1.733				
$DS_{Ag5}^S$	1	1/2/0.0015	<b>0.015</b>	0.381	1.651	1/2/0.0007	<b>0.044</b>	0.229	1.240
	2	1/4/0.0015	0.018	0.380	1.640	1/4/0.0007	0.045	0.216	1.224
	3	2/4/0.0015	0.018	0.378	1.636	2/4/0.0007	<b>0.044</b>	0.222	1.233
	4	2/4/0.0015	0.017	0.378	1.638	2/4/0.0007	0.047	0.223	1.235
$DS_{Ag7}^S$	1	1/1/0.0013	0.014	0.371	1.619	1/1/0.00062	<b>0.043</b>	0.231	1.244
	2	1/1/0.0013	0.013	0.372	1.628	1/1/0.0006	0.045	0.229	1.236
	3	2/4/0.00135	0.014	0.377	1.628	2/4/0.0006	0.050	0.218	1.222
	4	1/2/0.00135	<b>0.011</b>	0.376	1.637	1/2/0.0006	0.048	0.226	1.242
$DS_{Ag9}^S$	1	1/2/0.00115	<b>0.024</b>	0.413	1.648	1/2/0.0005	0.065	0.252	1.230
	2	1/2/0.00115	0.028	0.412	1.641	1/2/0.0005	<b>0.063</b>	0.254	1.234
	3	1/2/0.00115	0.027	0.411	1.642	1/2/0.0005	0.067	0.252	1.226
	4	1/4/0.0012	0.026	0.395	1.650	1/4/0.0005	<b>0.063</b>	0.250	1.240
$DS_{Ag11}^S$	1	1/1/0.0011	<b>0.016</b>	0.395	1.646	1/1/0.0005	0.056	0.232	1.236
	2	1/2/0.00115	0.017	0.403	1.656	1/2/0.0005	0.052	0.235	1.240
	3	1/3/0.00115	0.018	0.395	1.644	1/3/0.0005	<b>0.050</b>	0.228	1.236
	4	1/2/0.0011	0.017	0.399	1.653	1/2/0.0005	0.055	0.245	1.250

coalitions gives the best results. As was previously stated, for ensemble of three classifiers the use of weights have not changed the quality of classification.

## 7.2 Results with the Vehicle Silhouettes Data Set

The results of the experiments with the Vehicle Silhouettes data set are presented in Tables 22, 23, 24. Each table illustrates approach that was used to generate the coalitions: *Approach 1, 2 and 3*. The information in the tables is similar to the previous tables.

In Tables 22, 23, 24 the best results were in bold.

Based on the results given in Tables 22, 23 and 24 it can be concluded that for the Vehicle Silhouettes data set, there were no such significant differences between the results obtained using Approaches 1, 2 and 3 as it was in the case of the Soybean data set. One can certainly be said that Approaches 2 and 3 are better than Approach 1. Note that, Approach 3 generates only slightly better results than Approach 2. Comparing the methods of determining the strength of the coalition, most often, the

**Table 28** Comparison the methods of determining the strength of the coalition (Approach 3 and Landsat Satellites)

System	Weights	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$
$DS_{Ag3}^S$	1	1/4/0.0014	0.023	0.406	1.768				
	2	1/4/0.0014	0.022	0.402	1.751				
	3	1/4/0.0014	<b>0.021</b>	0.402	1.754				
	4	1/4/0.00135	0.022	0.396	1.721				
$DS_{Ag5}^S$	1	5/2/0.00147	<b>0.013</b>	0.372	1.643	5/2/0.0007	<b>0.042</b>	0.230	1.241
	2	2/2/0.0015	0.014	0.378	1.632	2/2/0.0007	0.045	0.224	1.232
	3	3/2/0.00155	<b>0.013</b>	0.384	1.653	3/2/0.0007	0.044	0.224	1.230
	4	4/5/0.0015	0.017	0.376	1.643	4/5/0.0007	0.045	0.220	1.232
$DS_{Ag7}^S$	1	1/1/0.0014	0.012	0.383	1.655	1/1/0.0007	0.045	0.227	1.244
	2	1/1/0.0014	0.014	0.378	1.653	1/1/0.0007	<b>0.042</b>	0.225	1.242
	3	1/1/0.0014	0.013	0.377	1.636	1/1/0.0007	0.043	0.225	1.242
	4	1/2/0.00135	<b>0.011</b>	0.385	1.640	1/2/0.0006	0.051	0.230	1.236
$DS_{Ag9}^S$	1	1/2/0.00125	<b>0.017</b>	0.392	1.630	1/2/0.0006	<b>0.050</b>	0.240	1.244
	2	1/5/0.0014	<b>0.017</b>	0.387	1.639	1/5/0.0007	0.051	0.234	1.242
	3	3/5/0.0013	0.023	0.407	1.667	3/5/0.0006	<b>0.050</b>	0.247	1.258
	4	1/2/0.0012	0.021	0.393	1.633	1/2/0.0006	0.063	0.257	1.253
$DS_{Ag11}^S$	1	1/1/0.00125	0.014	0.387	1.632	1/1/0.0006	<b>0.044</b>	0.224	1.245
	2	1/1/0.0013	<b>0.013</b>	0.387	1.627	1/1/0.0006	0.046	0.218	1.229
	3	1/1/0.0013	0.014	0.389	1.648	1/1/0.0006	<b>0.044</b>	0.222	1.239
	4	1/1/0.0012	0.018	0.384	1.622	1/1/0.0005	0.053	0.218	1.220

best results were generated by the size of the coalition method and the unambiguous of the decisions made by the classifiers method.

Table 25 shows the results obtained without the use of coalitions' weights. The best results within this table were in bold. Comparing results that were presented in Tables 22–25, it can be said that much better results were generated when the strength of the coalition are taken into account. When there is a larger number of classifiers we can expect a greater improvement. For three classifiers this improvement is negligible, or as for Approach 3, even worse results were obtained. This situation seems to be logical and it was already been commented earlier.

### 7.3 Results with the Landsat Satellites Data Set

The results of the experiments with the Landsat Satellites data set are presented in Tables 26, 27, 28. Each table illustrates approach that was used to generate the coalitions: Approach 1, 2 and 3. The information in the tables is similar to the previous tables.

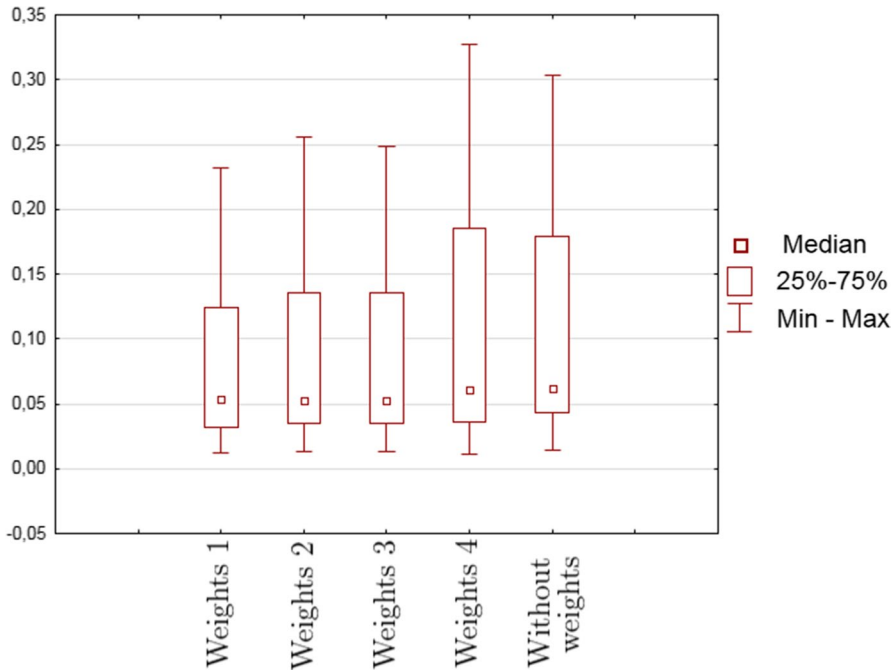
In Tables 26, 27, 28 the best results were in bold.

For the Landsat Satellites data set, the best results were generated using Approach 3, while the worst for Approach 1. Differences in the quality of

**Table 29** Comparison of the efficiency of inference using Approaches 1, 2, 3 and without the strength of the coalition (Landsat Satellites)

System	$m_1/m_2/\epsilon$	$e$	$e_{ONE}$	$\bar{d}_{DS}$	$t$
Approach 1					
$DS_{Ag3}$	1/3/0.0013	0.041	0.409	1.758	4.47
$DS_{Ag5}$	1/1/0.0014	<b>0.018</b>	0.360	1.607	5.19
	1/1/0.0007	<b>0.045</b>	0.235	1.245	5.19
$DS_{Ag7}$	3/5/0.0013	<b>0.014</b>	0.366	1.596	20.34
	3/5/0.0006	0.051	0.224	1.241	20.34
$DS_{Ag9}$	2/50.0012	<b>0.042</b>	0.400	1.618	42.58
	2/5/0.0006	0.077	0.263	1.253	42.58
$DS_{Ag11}$	1/3/0.00153	0.041	0.376	1.632	38.55
	1/3/0.0009	0.071	0.264	1.259	38.55
Approach 2					
$DS_{Ag3}$	1/3/0.0013	0.038	0.401	1.720	4.49
$DS_{Ag5}$	4/3/0.00154	0.019	0.379	1.627	6.41
	4/3/0.0008	0.048	0.238	1.258	6.41
$DS_{Ag7}$	1/5/0.0015	0.027	0.375	1.611	18.06
	1/5/0.0008	<b>0.043</b>	0.231	1.243	18.06
$DS_{Ag9}$	1/4/0.0015	0.043	0.378	1.606	27.49
	1/4/0.0008	<b>0.066</b>	0.256	1.239	27.49
$DS_{Ag11}$	1/3/0.00157	<b>0.039</b>	0.387	1.622	34.09
	1/3/0.0008	<b>0.062</b>	0.245	1.233	34.09
Approach 3					
$DS_{Ag3}$	1/3/0.0015	<b>0.035</b>	0.388	1.711	5.02
$DS_{Ag5}$	2/2/0.0018	0.028	0.382	1.607	5.59
	2/2/0.001	0.047	0.240	1.245	5.59
$DS_{Ag7}$	1/3/0.0021	0.031	0.371	1.612	7.39
	1/3/0.0013	0.049	0.246	1.240	7.39
$DS_{Ag9}$	1/1/0.0027	0.043	0.389	1.613	6.56
	1/1/0.0015	0.071	0.253	1.237	6.56
$DS_{Ag11}$	1/1/0.0032	0.041	0.377	1.635	7.45
	1/1/0.0019	0.063	0.244	1.245	7.45

classification are noticeable, especially for a greater number of classifiers. Thus, using the weights, it was possible to translate better ability to recognize the relationship between the classifiers of Approach 3 (which was already been demonstrated in paper Przybyła-Kasperek 2017) on the quality of classification. It should be noted that for Approach 1, the method that was used to calculate weights of coalitions is irrelevant. For all four methods the same results were obtained. This is due to poor quality of coalitions that were built using Approach 1. It can not be stated that one of the methods of calculating the weights of coalitions is the best. The only thing that can be said is that the method considering the unambiguous of the decisions made by the coalition usually generates the worst results.



**Fig. 4** Comparison of the classification error  $e$  for four different methods of determining coalitions' strength and for approach without any weights

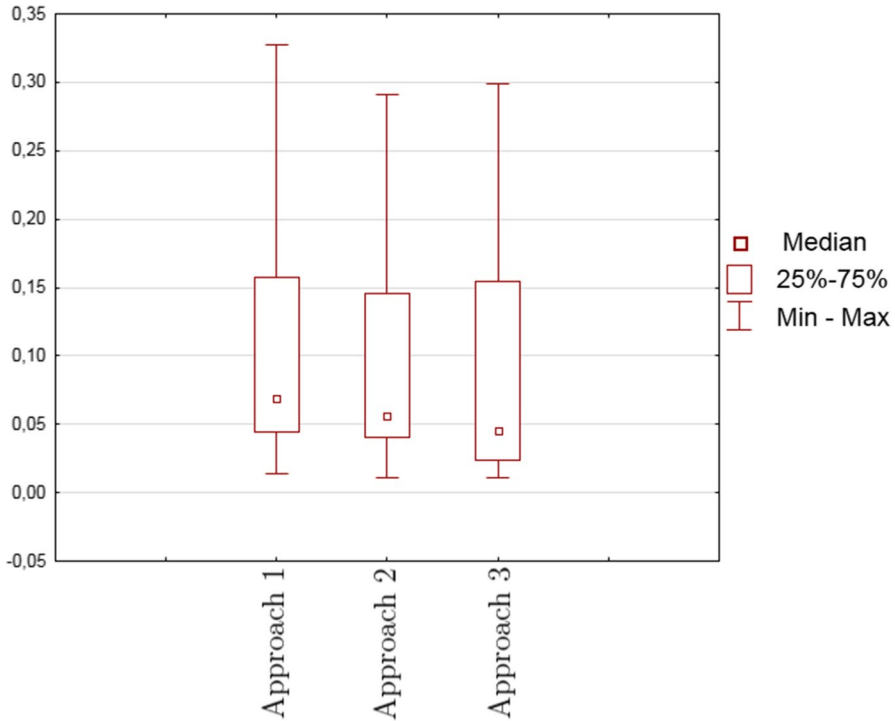
Comparison of results obtained without the use of the strength of the coalition was also performed for the Landsat Satellites data set. Results are presented in Table 29. The best results within this table were in bold. It has been noted that for Approaches 1 and 2 not always the use of weights improves the quality of classification. However, for a greater number of classifiers improving is noticeable. Only for Approach 3 we can always be sure of improving the results when using weights, this improvement is very significant and noticeable.

#### 7.4 Comparison of Methods of Coalitions' Generation and Methods of Determining Coalitions' Strength

Based on the results described in the previous sections, it was found that the use of coalitions' strength has improved the quality of classification, regardless of the approach to creating coalitions that was used. It is difficult to designate which of the ways of calculating weights is the best. However, the results show that the last method—the unambiguous of the decisions made by the coalition—achieves the worst results. The first method—the size of the coalition—very often produces good results. A statistical analysis was performed to confirm this observations.

The results were divided into five groups—four groups for different methods of calculating weights and one group consists of the values of error obtained without





**Fig. 5** Comparison of the classification error  $e$  for three different approaches to generating coalitions

calculating the weights. The Friedman's test was performed at first. The test confirmed that the differences between the classification error in these five groups are significant with a level of  $p = 0.000001$ . Then, in order to determine the pairs of methods between which statistically significant differences occur, a nonparametric Wilcoxon each pair test and a parametric  $t$ -test for dependent groups were performed. Both tests shown that there is no significant difference between pairs of methods 2, 3 (the unambiguous of the decisions made by the classifiers and the size of the coalition and the unambiguous of the taken decisions) and methods 4, 5 (the unambiguous of the decisions made by the coalition and without calculating the weights). For all other pairs the importance of differences was confirmed at a significance level less than 0.05.

The box-whiskers chart which presents statistics for these five groups was created (Fig. 4). It can be clearly seen that the box for the first method for calculating the weights of the coalitions is located the lowest in the graph. The second and the third methods have comparable boxes, and the boxes for the last method of calculating weights and approach without weights are the largest and are at the top of the graph.

The final conclusion is that the first method is the best for use. The second and the third methods are in the second position, and it does not make sense to use the fourth method.

Similar tests were made to find out which of the approaches to coalition creation is the best. In the previous sections it was stated that Approach 3 is the best, and Approach 1 is the worst. In order to statistical justification of this, all results are divided into three groups, each representing different approach. The Friedman's test confirmed that the differences between the classification error for these three approaches are significant with a level of  $p = 0.000001$ . Both tests—a nonparametric Wilcoxon each pair test and a parametric  $t$ -test for dependent groups—confirmed that for all pairs of different approaches the differences are statistically significant at a significance level less than 0.0004.

The box–whiskers chart that presents statistics for these three groups is shown on Fig. 5. The smallest median of error values is observed for Approach 3 and the largest for Approach 1. The final conclusion is that the best way of creating coalitions is Approach 3.

## 8 Summary

In this paper, three approaches to use the Pawlak's conflict model in order to analyze relations between classifiers and to creation coalitions of classifiers were considered. Four methods of calculating coalitions weights, in order to take advantage of the generated coalitions, were proposed. The results of the experiments on three data sets that were dispersed in five different versions were presented. The obtained results were compared and hypotheses were statistically justified. It has been found that Approach 3 to coalition creation generates the best results. In terms of the weights of the coalitions the method that is based on the size of the coalition is the best. A little worse results were generated by the unambiguous of the decisions made by the classifiers method and the size of the coalition and the unambiguous of the taken decisions method.

In this paper it was shown that the use of coalitions' weights improves the quality of classification as long as the coalitions are correctly identified. Three approaches of using Pawlak's analysis model were used to determine the coalitions, but only one of them has the ability to generate well-formed coalitions. Further modifications of application Pawlak's conflict analysis method to a dispersed system are planned in the future work. It is planned to use the approaches proposed in Lang et al. (2017); Yao (2019) to develop ideas presented in this paper.

**Acknowledgements** The author thanks editors and reviewers for their critical and constructive comments.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bloch I (1996) Fuzzy relative position between objects in images: a morphological approach. *ICIP* 2:987–990
- Deja R (2002) Conflict analysis. *Int J Intell Syst* 17(2):235–253. <https://doi.org/10.1002/int.10019>
- Fatima S, Wooldridge M, Jennings N (2005) A comparative study of game theoretic and evolutionary models of bargaining for software agents. *Artif Intell Rev* 23:185–203
- Kersten G, Lai H (2007) Negotiation support and e-negotiation systems: an overview. *Group Decis Negot* 16:553–586
- Lang GM, Miao DQ, Cai MJ (2017) Three-way decision approaches to conflict analysis using decision-theoretic rough set theory. *Inf Sci* 406–407:185–207
- Lopes F, Novais AQ, Mamede N, Coelho H (2008) Negotiation among autonomous agents: experimental evaluation of integrative strategies. *Artif Intell Rev* 29:1–44
- Panov P, Džeroski S (2007) Combining bagging and random subspaces to create better ensembles. In: Berthold RM, Shawe-Taylor J., Lavra N (eds) *Advances in intelligent data analysis VII. IDA 2007*. LNCS, vol 4723. Springer, Berlin
- Pawlak Z (2006) Conflicts and negotiations. In: Wang G-Y, Peters JF, Skowron A, Yao Y (eds) *RSKT 2006*. LNCS (LNAD), vol. 4062, Springer, Heidelberg, pp 12–27. [https://doi.org/10.1007/11795131\\_2](https://doi.org/10.1007/11795131_2)
- Pawlak Z (1984) On conflicts. *Int J Man-Mach Stud* 21:127–134
- Pawlak Z (2005) Some remarks on conflict analysis. *Eur J Oper Res* 166:649–654. <https://doi.org/10.1016/j.ejor.2003.09.038>
- Polikar R, Parikh D, Mandayam S (2006) Multiple classifier systems for multisensor data fusion. In: *Proceedings of the 2006 IEEE sensors applications symposium, 2006*. IEEE
- Przybyła-Kasperek M (2017) Directions of use of the Pawlak’s approach to conflict analysis. In: Wang G, Skowron A, Yao Y, Ślęzak D, Polkowski L (eds) *Thriving rough sets: 10th anniversary—honoring Professor Zdzisław Pawlak’s Life and Legacy and 35 Years of Rough Sets*. Springer, Berlin, 329–351
- Przybyła-Kasperek M (2017) Methods based on Pawlak’s model of conflict analysis—medical applications, rough sets. In: Polkowski L, Yao Y, Artiemjew P, Ciucci D, Liu D, Ślęzak D, Zielosko B (eds) *International joint conference, IJCRS 2017, Olsztyn, Poland, July 3–7, 2017, proceedings, Part I*, LNCS, 249–262
- Przybyła-Kasperek M, Wakulicz-Deja A (2014) Global decision-making system with dynamically generated clusters. *Inf Sci* 270:172–191
- Przybyła-Kasperek M, Wakulicz-Deja A (2014) A dispersed decision-making system: the use of negotiations during the dynamic generation of a systems structure. *Inf Sci* 288:194–219
- Przybyła-Kasperek M, Wakulicz-Deja A (2016) Global decision-making in multi-agent decision-making system with dynamically generated disjoint clusters. *Appl Soft Comput* 40:603–615
- Rahwan I, Ramchurn S, Jennings N, McBurney P, Parsons S, Sonenberg L (2004) Argumentation-based negotiation. *Knowl Eng Rev* 18:343–375
- Ramanna S, Peters JF, Skowron A (2007) Approaches to conflict dynamics based on rough sets. *Fundam Inform* 75(1–4):453–468
- Ramanna S, Peters JF, Skowron A (2006) Generalized conflict and resolution model with approximation spaces. In: *Rough sets and current trends in computing, 5th international conference, RSCTC 2006, Kobe, Japan, Nov 6–8, 2006, proceedings*, 274–283. [https://doi.org/10.1007/11908029\\_30](https://doi.org/10.1007/11908029_30)
- Shoemaker L, Banfield RE, Hall LO, Bowyer KW, Kegelmeyer WP (2008) Using classifier ensembles to label spatially disjoint data. *Inf Fusion* 9(1):120–133
- Skowron A, Deja R (2002) On some conflict models and conflict resolutions. *Romanian J Inf Sci Technol* 3(1–2):69–82
- Skowron A, Ramanna S, Peters JF (2006) Conflict analysis and information systems: a rough set approach. In: *Rough sets and knowledge technology, first international conference, RSKT 2006, Chongqing, China, July 24–26, 2006, proceedings*, 233–240. [https://doi.org/10.1007/11795131\\_34](https://doi.org/10.1007/11795131_34)
- Stefanowski J (2005) An experimental study of methods combining multiple classifiers-diversified both by feature selection and bootstrap sampling. In: Atanassov TK, Kacprzyk J, Krawczak M, Szmidt E (eds) *Issues in the representation and processing of uncertain and imprecise information*. Akademicka Oficyna Wydawnicza EXIT, Warsaw, pp 337–354
- Tang EK, Suganthan PN, Yao X (2006) An analysis of diversity measures. *Mach Learn* 65(1):247–271

Yao YY (2019) Three-way conflict analysis: reformulations and extensions of the Pawlak model. *Knowledge-Based Systems* 180:26–37

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.