Information Sciences 270 (2014) 172–191

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Global decision-making system with dynamically generated clusters

Małgorzata Przybyła-Kasperek *, Alicja Wakulicz-Deja

University of Silesia, Institute of Computer Science, Będzińska 39, 41-200 Sosnowiec, Poland

ARTICLE INFO

Article history:
Received 21 February 2013
Received in revised form 29 October 2013
Accepted 9 February 2014
Available online 22 February 2014

Keywords:
Decision-making system
Global decision
Relation of friendship
Relation of conflict
Coalition

ABSTRACT

This paper discusses the issues related to the process of global decision-making on the basis of knowledge which is stored in a dispersed form (several local knowledge bases or classifiers). In the paper a decision-making system is described. In this system, the classification process of the test object can be divided into several steps. In the first step, we investigate how particular classifiers classify a test object. We describe this using probability vectors over decision classes. We cluster classifiers with respect to similarities of the probability vectors. For every cluster, we find a kind of combined information. Finally, we classify the given test object by voting among clusters, using the combined information from each of clusters.

The paper proposes a new approach to the organization of the structure of a decision-making system, which operates on the basis of dispersed knowledge. In the presented system, the classifiers are combined into groups called clusters in a dynamic way. We seek to designate groups of classifiers that classify the test object in a similar manner. The groups of classifiers are not disjoint sets. We use overlapping clusters because this is a more suitable representation of classification compatibility. It is assumed that, if the classifier classifies the test object in an ambiguous way, it should belong to several clusters. Then, a process of the elimination of inconsistencies in the knowledge is implemented in the created groups. Global decisions are made by using one of the methods for the analysis of conflicts.

1. Introduction

The problem of making decisions on the basis of dispersed knowledge stored in many local knowledge bases is examined in this paper. This problem concerns complex knowledge bases in which the possibility of the cooperation of local bases in order to reach a common decision (a global decision) is very important. In this paper a new approach to the organization of a system structure that uses dispersed knowledge is proposed.

In the proposed approach, we assume that each local knowledge base is managed by one agent. This situation can be seen as a set of classifiers, where each of the classifiers has access to a different knowledge set. In order to take a global decision, agents are combined into groups. In the new approach described in this paper, groups are created in a dynamic way. In earlier papers [22–24,30–33], a multi-agent system was also considered, but that system has a different structure. In the approaches used previously, the static structure of the system or a dynamic structure of the multi-agent system in which

* Corresponding author. Tel.: +48 32 3689757.
E-mail address: malgorzata.przybyla-kasperek@us.edu.pl (M. Przybyła-Kasperek).

http://dx.doi.org/10.1016/j.ins.2014.02.076
0020-0255/

groups of agents are disjoint were considered. In the system proposed in this paper, a dynamic structure in which groups of agents are not disjoint sets is considered. In the proposed system, we aim to designate homogeneous groups of agents. The agents which agree on the classification for a test object into the decision classes are combined in a group. Many real life applications are characterized by situations in which overlapping clusters would be a more suitable representation. Very often agents that take part in the negotiations are unable to take one, explicit decision. Sometimes a few decisions are acceptable to one decision unit. In this situation agents should belong to several clusters. From a more technical point of view, in order to define clusters for each agent, a probability vector over decision classes is determined. This vector describes the classification of a test object that is made by the agent. Then agents are clustered with respect to the similarities of probability vectors. In order to identify groups of agents, the concepts of a friendship relation and a conflict relation, which were introduced by Pawlak in the papers [16–18], are used.

In the paper, a multi-agent system with a hierarchical structure is used. Local decisions are taken based on the knowledge of agents from one group. For every cluster, a kind of combined information is determined. Since the sets of attributes, conditions on the basis of which agents classify the test object do not have to be disjoint, an inconsistency in knowledge can occur. Therefore, a method for the elimination of inconsistencies in the knowledge is discussed here. Finally, the test object is classified by voting among clusters, using the combined information from each of the clusters. The problem of conflict analysis arises because the inference is being conducted in groups of knowledge bases. By a conflict, we mean a situation in which conflicting decisions are taken for the specified set of conditions on the basis of the knowledge that is stored in different groups of knowledge bases. This paper discusses two methods of conflict analysis (proposed in earlier papers [31–33]) that allow inference in spite of the presence of conflicts.

The paper is organized as follows. The second section presents the related papers. The third section describes the structure of a decision-making system. This section is divided into three parts. The first part of this section contains a high-level design description of the systems structure and the second part contains more technical issues. The last part of the third section presents an example of creating clusters. The fourth section describes the method of the elimination of inconsistencies in the knowledge. The fifth section describes the method of conflict analysis. The sixth section shows a description and the results of experiments carried out using some data sets from the UCI repository. The article concludes with a short summary in the seventh section.

2. Related work

The main aim of this paper is to propose a system in which knowledge bases will be combined into groups or coalitions in a dynamic way. The theory of negotiations and the formation of coalitions is an important issue of social interaction and it is studied in various branches of the social sciences as well as in computer science. A brief overview of various negotiation models that have been proposed in the literature can be found in the paper [14]. Zeng and Sycara [34] proposed a sequential decision-making model, called Bazaar, which is able to learn. Faratin et al. [7] presented a model for bilateral service-oriented negotiation that defines a range of strategies and three groups of concession tactics. Lopes et al. [13] presented a negotiation model that formalizes various problem-solving and concession strategies. Nguyen and Jennings [15] presented a model that handles one-to-many negotiations in service-oriented contexts.

In the papers of Pawlak [16–18], a different approach to the issue of coalition formation was considered. This model describes a conflict situation in which the agents have decided to analyze the conflict by using a peaceful method. In such a situation the relations of conflict, friendship and neutrality were defined and the method of formulating the coalition, as well as the method of calculating the intensity of the conflict were proposed. In this paper, some issues of conflict analysis and coalition formation that were given in Pawlak's model are used.

Conflict analysis has its roots in game theory where two of the basic methods used to measure the power of agents is the Shapley–Shubik power index and the Banzhaf power index. In the paper [21] of Polkowski and Araszkiewicz the Shapley–Shubik power index was applied in order to analyze a game on the basis of partial data about the game. A number of attempts have been made to solve the conflict. As was mentioned earlier, one of the solutions concerning the analysis of conflicts was suggested by Pawlak [16,18]. Other authors have also investigated this issue [4,25]. In the paper of Szczuka and Ślęzak [28] a generalized neural network model is used to learn voting between local classifiers from data. A very interesting approach to resolve conflicts was proposed by Bazan [1]. In this method, it is assumed that the set of classifiers is given. The conflicts that arose between the classifiers are solved by the construction and analysis of a special table. This special table is called a conflict table and it is constructed on the basis of the classification results of objects from the validation table. This approach will be considered in a future work.

The issue of making decisions based on distributed knowledge is widely considered in the literature. For example, this issue is considered in the multiple model approach [8,11,12]. In a multiple classifier system, an ensemble is constructed on the basis of base classifiers. The aim of this approach is to reduce the misclassification at the cost of increased computational complexity. In the papers [2,3,9,10], another (different from that which is considered in this paper) approach to the problem of classification on the basis of several different decision tables can be found. In methods of Distributed Data Mining (DDM), it is assumed that the data are collected and stored in different decision tables representing either horizontally or vertically partitioned. In the paper [29] an approach in which the goal is to learn an ensemble of relatively weak classifiers based on some subsets of attributes, which together provide the correct answers using a very simple model of conflict