

Aggregation Procedure Based on Majority Principle for Collective Identification of Firm's Crucial Knowledge

Inès Saad^{1,2} and Salem Chakhar³

¹ MIS, University of Picardie Jules Vernes, 33 Rue Saint Leu, 80039 Amiens Cedex 01, France

² Amiens School of Management, 18 Place Saint-Michel, 80000 Amiens, France

³ LAMSADE, University Paris-Dauphine

ines.saad@u-picardie.fr, salem.chakhar@dauphine.fr

Abstract. It is very important to identify, preserve, and transfer knowledge to those who need it within firm. However, the identification of knowledge and especially tacit knowledge is a complex process because knowledge cannot be measured quantitatively. In this paper we present an approach for inducing a set of collective decision rules representing a generalized description of the preferential information of a group of decision makers involved in a multicriteria classification problem to identify crucial knowledge to be capitalized.

Keywords: Knowledge management, Knowledge identification, DRSA, Aggregation procedure.

1 Introduction

Firms should invest in knowledge management projects in order to capitalize valuable knowledge. Researchers in knowledge engineering and knowledge management have been focusing on the problems of acquisition, preservation and transfer of knowledge [4]. However, considering the large amount of knowledge to be capitalized, the firms must focalize on only the valuable knowledge. Not enough works exist concerning the identification of knowledge on which it is necessary to capitalize [2][3][6] [7], thus, we have proposed a multicriteria method based on DRSA (Dominance-based Rough Set Approach) [5] to justify a situation where knowledge capitalization is advisable. The methodology is composed of three phases. The first phase is relative to constructive learning. The second phase consists on inferring the preference model of the decision makers, and finally the third phase applies the previous model to evaluate other knowledge used in the same project or other similar projects. Using our method [6] allows taking into account the preferences of decision makers that can be different or even contradictory while exploiting and managing their multiple point of views to evaluate knowledge, without using a quantitative measuring approach. Moreover, because of the large amount of knowledge to analyze, the large number of decision makers involved in the assignments of knowledge, contradictory opinions that decision makers can have, it is necessary to solve conflicts between decision makers. In fact, the methodology requires, when inconsistency situations are identified, that the analyst conducts an in-depth discussion with the different decision makers in order to solve the conflicts [1]. This is a time-consuming and difficult task.

The aim of this paper is to propose an aggregation procedure based on the majority principle which is defined through the concordance and discordance powers.

The rest of the paper is structured as follows. Section 2 presents the background. Section 3 presents our aggregation procedure. The application of the method and results are presented in section 4. Section 5 summarizes our contribution.

2 Background

The DRSA [5] is a rough sets-based multicriteria classification method. This method has been developed to overcome the shortcomings of rough set in multicriteria classification problems. Indeed, conventional rough set theory based on indiscernibility cannot be used to deal with multicriteria classification problems since attributes are preference ordered. The basic idea of DRSA is to replace indiscernibility principle with dominance.

Information about decision objects are often represented in terms of an *information table* where rows correspond to *objects* and columns correspond to *attributes* and entries are *attributes-values*. Formally, the information table S is a 4-tuple $\langle U, Q, V, f \rangle$ where:

- U is a finite set of objects,
- Q is a finite set of attributes,
- $V = \bigcup_{q \in Q} Vq$ Vq is a domain of the attribute q , and
- $F: U \times Q \rightarrow V$ an information function defined such that $f(x, q) \in Vq, \forall q \in Q, \forall x \in U$.

The set of attributes Q is often divided into two sub-sets: sub-set C of *condition attributes* and sub-set D of *decision attributes*. In this case, the information table is called *decision table*.

For the purpose of this paper, a series of assumptions, which are appropriate in multicriteria classification problems, are established. The domain (or scale) of condition attributes are supposed to be ordered to decreasing or increasing preference. Such attributes are called *criteria*. Without loss of generality, we assume that the preference is increasing with value of $f(., q)$ for every $q \in C$. We also assume that the set of decision attributes D is a singleton $\{d\}$. Decision attribute d makes a partition of U into a finite number of decision classes $Cl = \{Cl_t, t \in T\}, T = \{0, \dots, n\}$, such that each $x \in U$ belongs to one and only one class in Cl . Further, we suppose that the classes are preference-ordered, i.e. for all $r, s \in T$, such that $r > s$, the objects from Cl_r are preferred to the objects from Cl_s .

The basic idea of rough set approach is the approximation of knowledge generated by the decision attributes by “*granules of knowledge*” generated by condition attributes. The sets to be approximated are defined as follows:

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s, t = 0, \dots, n.$$

The set Cl_t^{\geq} is called the upward union. The assertion $x \in Cl_t^{\geq}$ means that “*x belongs to at least class Cl_t* ”. The set Cl_t^{\leq} is called the downward union. The assertion $x \in Cl_t^{\leq}$ means that “*x belongs to at most Cl_t* ”.

The P-lower approximation of Cl_t^{\geq} contains all the objects with P-dominating set are assigned with certitude to classes at most as good as Cl_t . The P-upper approximation of Cl_t^{\geq} contains all the objects with P-dominating set are assigned to a class at least as good as Cl_t . We can also define the P-boundary sets of Cl_t^{\geq} $Bnp(Cl_t^{\geq})$ contains all the objects which are assigned both to a class better than Cl_t and to one or several classes worse than Cl_t .

The quality of classification is defined by the following ratio:

$$\gamma_P = \frac{card(U - (\bigcup_{t=1, \dots, n} Bn_p(Cl_t^{\geq})))}{card(U)} \tag{1}$$

The ratio expresses the percentage of objects that are assigned with certitude in a given class.

3 Collective Decision Rules Construction Approach

The proposed approach is composed of three phases: (i) individual classification, (ii) aggregation, and (iii) generation of collective decision rules. The main input of the approach is a common information table with a finite set of knowledge to be evaluated and a finite set of criteria $C = \{1, \dots, q, \dots, m\}$. The output is a collection of collective decision rules representing a generalized description of the preferential information provided by the different decision makers.

Let $DM = \{DM_1, \dots, DM_i, \dots, DM_h\}$ be a finite set of decision makers corresponding to h decision attributes $D_1, \dots, D_i, \dots, D_h$. Further, we suppose that decision attributes are defined on the same domain. In addition, we suppose that each decision maker $i \in H$ has a preference order on the universe U and that this preference order is represented by a finite set of preference ordered classes:

$$Cl_i = \{Cl_{t,i}, t \in T_i\}, T_i = \{0, \dots, n_i\},$$

3.1 Phase 1: Individual Classification

In this first phase, each decision maker uses the common information table I to construct its own decision table S_i defined as $\langle U, CUD_i, V, f_i \rangle$ where, for each $i \in H$, D_i is a new decision attribute and f_i is the information function, both associated with decision maker i .

Then, each decision maker $i \in H$, with the help of the analyst, runs the DRSA method using its decision table S_i as input. At this level, DRSA does not need any extension, because each decision maker uses its own preference information.

In terms of this phase, the classification conducted by each decision maker $i \in H$ is characterized, among others, by:

- the P-lower approximation and P-boundary of $Cl_{t,i}^{\leq}$ and $Cl_{t,i}^{\geq}$, for each $t \in Ti$, and
- the quality of classification γ_P^i defined as :

$$\gamma_P^i = \frac{\text{card}(U - (\bigcup_{t=1, \dots, ni} Bn_p(Cl_{t,i}^{\geq})))}{\text{card}(U)} \tag{2}$$

These information represent the inputs for the next phase.

3.2 Phase 2: Aggregation Procedure

The aggregation procedure is implemented in terms of three steps: (i) normalization, (ii) computing the concordance and discordance powers and (iii) definition of assignment.

3.2.1 Step 2.1: Normalization

The objective of this first step is to standardize the quality of classification γ_P^i ($i \in H$) using the following formula:

$${}^i\gamma'_P = \frac{1}{h} \sum_{i=1}^h \gamma_P^i, (i \in H). \tag{3}$$

It is clear that $\sum_{i=1}^h {}^i\gamma'_P = 1$

3.2.2 Step 2.2: Computing the Concordance and Discordance Powers

The aggregation procedure is based on the majority principle which is defined through the concordance and discordance powers. These powers are computed based on the outputs of individual classification phase (i.e. P-lower approximation, P-boundary and quality of classifications).

Concordance Power

Based on the lower approximations of $Cl_{t,i}^{\geq}$, we define, for each knowledge $x \in U$ and a class $Cl_t \in Cl$, the sets $L(x, Cl_t^{\geq})$ as follows :

- $L(x, Cl_t^{\geq}) = \{i: i \in H \wedge x \in \underline{P}(Cl_{t,i}^{\leq})\} (t=1, \dots, n)$,

The first set represents the decision makers for which knowledge x belongs to the lower approximation of Cl_t^{\leq} . The second one represents the decision makers for which knowledge x belongs to the lower approximation of Cl_t^{\geq} . Next, the concordance powers for the assignment of x to Cl_t^{\geq} are computed as follows:

$$L^+(x, Cl_t^{\geq}) = \sum_{i \in L(x, Cl_t^{\geq})}^h i \gamma_p', (t = 0, \dots, n - 1) \tag{4}$$

The number $L^+(x, Cl_t^{\geq}) \in [0, 1]$ measures the power of coalition of decision makers that assign x to the lower approximation of Cl_t^{\geq} .

Discordance Power

Based on the lower approximations of $Cl_{t,i}^{\geq}$, we define, for each knowledge $x \in U$ and a class $Cl_t \in Cl$, the set $B(x, Cl_t^{\geq})$ as follows :

- $B(x, Cl_t^{\geq}) = \{i: i \in H \wedge x \in Bnp(Cl_{t,i}^{\geq})\} (t=1, \dots, n)$,

The first set represents the decision makers for which knowledge x belongs to the lower approximation of Cl_t^{\leq} . The second one represents the decision makers for which knowledge x belongs to the lower approximation of Cl_t^{\geq} . Next, the concordance powers for the assignment of x to Cl_t^{\leq} and to Cl_t^{\geq} are computed as follows:

$$B^+(x, Cl_t^{\geq}) = \sum_{i \in B(x, Cl_t^{\geq})}^h i \gamma_p', (t = 0, \dots, n - 1) \tag{5}$$

The number $B^+(x, Cl_t^{\geq}) \in [0, 1]$ measures the power of coalition of decision makers that assign x to the lower approximation of Cl_t^{\geq} .

3.2.3 Step 2.3: Definition of Assignment

Let $\theta \in [.5, 1.0]$ be a majority threshold and $\theta' \in [.5, 1.0]$ be a veto threshold. Based on the concordance and discordance powers, we may distinguish four situations for the assignment of x to Cl_t^{\geq} :

	$B^+(x, Cl_t^{\geq}) < \theta'$	$B^+(x, Cl_t^{\geq}) \geq \theta$
$L^+(x, Cl_t^{\geq}) \geq \theta$	$x \in Cl_t^{\geq}$	$x \notin Cl_t^{\geq}$
$L^+(x, Cl_t^{\geq}) < \theta$	$x \notin Cl_t^{\geq}$	$x \notin Cl_t^{\geq}$

These situations can be summarized by the following assignment rule:

Rule: If $L^+(x, Cl_t^{\geq}) \geq \theta \wedge B^+(x, Cl_t^{\geq}) < \theta'$, then $x \in Cl_t^{\geq}$ else $x \notin Cl_t^{\geq}$

This assignment rule can be explained as follows. A knowledge x is assigned to Cl_t^{\geq} if and only if:

- There is a “sufficient” majority of decision makers (in terms of their quality of classification) that assign x to Cl_i^{\geq}
- When the first condition holds, none of the minority of decision makers shows an “important” opposition to the assignment of x to Cl_i^{\geq} .

Thus the assignment rule is used to determine a collective decision table.

3.3 Phase 3: Generation of Collective Decision Rules

The objective of this phase is to apply the DRSA method to the collective decision table which was determined by the collective assignment of each knowledge to decision classes (crucial or not crucial). This decision table is used to infer a set of collective decision rules.

4 Conclusion

This paper details the issue of identification and evaluation of crucial knowledge and proposes an approach for inducing a set of collective decision rules in order to provide a conflict resolution in the context of crucial knowledge classification. The aim of the proposed approach is to manage conflicts between decision makers by the aggregation procedure based on the majority principle which is defined through the concordance and discordance powers.

The proposed methodology is based on the DRSA method. The original DRSA is a single decision maker oriented method. The proposed methodology permits to extend DRSA to group multicriteria classification problems. The methodology is composed of three phases: individual classification, aggregation and generation of the collective decision rules.

References

1. Belton, V., Pictet, J.: A framework for group decision using a MCDA model: sharing, aggregation or comparing individual information. *Journal of Decision Systems* 6(3), 283–303 (1997)
2. Brigui-Chtioui, I., Saad, I.: A mediating algorithm for multicriteria collaborative knowledge classification. In: Lytras, M.D., Damiani, E., Carroll, J.M., Tennyson, R.D., Avison, D., Naeve, A., Dale, A., Lefrere, P., Tan, F., Sipiør, J., Vossen, G. (eds.) *WSKS 2009. LNCS*, vol. 5736, pp. 147–155. Springer, Heidelberg (2009)
3. Brigui-Chtioui, I., Saad, I.: A multi-agent approach for collective decision making in knowledge management, *Group Decision and Negotiation* (to appear, 2010)
4. Dieng, R., Corby, O., Giboin, A., Rybière, M.: Methods and tools for corporate knowledge management. Technical report, INRIA, ACACIA project (1999)
5. Greco, S., Matarazzo, B., Slowinski, R.: Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research* 129, 1–47 (2001)

6. Saad, I., Rosenthal-Sabroux, C., Grundstein, M.: Improving the Decision Making Process in The Design project by Capitalizing on Company's Crucial Knowledge. *Group Decision and Negotiation* 14, 131–145 (2005)
7. Tseng, B., Huang, C.: Capitalizing on Knowledge: A Novel Approach to Crucial Knowledge Determination. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans* 35, 919–931 (2005)