

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Ad Hoc Networks Based on Rough Set Distance Learning Method

Yang Hui

College of Science, Xi'an University of Science and Technology, Xi'an, China

Abstract: With regard to a wireless sensor network, organizing sensors into a hierarchy of clusters with an objective of minimizing the total spent energy to communicate the information gathered by sensors to the information-processing center, rough set theory can serve as a useful pre-processing tool, removing superfluous attributes and suggesting validated rules and models on an objective and global basis, i.e., without any tuning for this particular application. Actually, data management for the WSN is very complex and there is wide scope for further research on intelligent data collection and dynamic reduction in this area so as to make the schemas more applicable. This study discuss the distance learning algorithm DLA (distance learning algorithm) to achieve, as well as it's the modern teaching of universal significance. Inductive learning is a research area in Artificial Intelligence. It has been used to model the knowledge of human experts by using a carefully chosen sample of expert decisions to infer decision rules. Rough set based inductive learning uses rough set.

Key words: Rough sets, distance learning algorithm, identification index, accuracy, nucleus

INTRODUCTION

A symbolic approach to achieve this aim is the rough set theory which identifies partial or total dependencies (i.e., cause-effect relations) in databases, eliminates redundant data and gives approach to null values, missing data, dynamic data and others (Wei *et al.*, 2010a). By using rough set theory, in wireless sensor networks, data aggregation algorithm operates on vast raw spatial-temporal redundancy information and extracts small amount of information useful for the final decision, thus, both communication overhead and energy consumption are significantly reduced.

As networks become increasingly popular course teaches the inherent mode makes online students and the lack of contact between educators and information feedback is increasingly prominent. In this study, we discuss algorithm implementation of distance education issues and demonstrates how to apply decision-making Tree to create a simple information sheet about through the realization of DLA (Pawlak, 1982; Wei *et al.*, 2008b). You can make the future of online learners understand the central part of the curriculum, teaching Education are also important parts of the course to make arrangements accordingly. Inductive Learning is a research area in Artificial Intelligence. It has been used to model the knowledge of human experts by using a carefully chosen sample of expert decisions to infer decision rules. Rough Set based Inductive Learning uses Rough Set.

Theory to compute decision rules. Guided online instruction is better. To solve the problem of lack of

contact between students and instructor, some rules on reasons for failure are needed. Such rules can advise new students about the rules which apply to them, based on past course grade history. If students pass, these rules nevertheless advise them on their weak areas, so they can better prepare for future courses because passing does not mean one hundred percent comprehension of the course material. If students fail, these rules inform them of the areas in which they are weak and suggest to them those sections they should focus on if they repeat the course.

Student specific information is not as useful to students because it does not tell them which concepts (sections) are needed as prerequisites for other concepts (sections), as the student could have had a bad day on the day of the quiz. In other words, the lack of contact problem can be solved by using inductive learning method to discover knowledge from the course grade history.

Rough Set theory, introduced by Pawlak (1991) and Wei *et al.* (2008b), is a mathematical tool for dealing with vagueness and uncertainty. Vagueness is caused by the ambiguity of exact meaning of terms used in a knowledge domain, uncertainty in data, or in knowledge itself. To deal with vagueness, normally statistics are used for handling likelihood. The advantage of rough set theory is that it does not need any preliminary or additional information about data (like probability in statistics, grade of membership, or the value of possibility in fuzzy set theory). Another advantage of the rough set approach is easy of use and its simple algorithms. RS is different from the general math set. The

uncertainty handling is derived from the points of knowledge. RS describes and analyzes the non-integration, the inaccuracy and the inconsistency effectively and then, the implicit knowledge and its laws are also discovered.

Pawlak (1982) and Wei *et al.* (2009) expound systematically RS theory. The basic RS concepts such as knowledge and non-distinction relation, lower and upper approximation, boundary region and rough membership function are given to illustrate. Born about only one decade or more, RS can practice greatly and many achievements are gained in many fields. RS has many distinguishing features, which is used in stocks data analysis, mode discerning, earthquake prediction, conflicts analysis, KDD, rough control and so forth. Much attention has been paid on RS development and achievements, which provides power for the uncertain information. Generally, both AI and its complicated information handling are based on classification, which is the basis of RS. It describes the classification as the specific space is divided. The equivalence relationship divided by RS is equal to that of knowledge. The knowledge conception in different scopes has different meanings. The space, U, expressed by discrete elements, is divided by the equivalence set, R and the knowledge is the results of U dividing by R. Therefore, the knowledge database can be defined as division of U by all the possible relation belonging to R, which is expressed as $K = (U, R)$. If the equivalence set has conflict with the data division, the process leading to uncertain division can be measured by roughness. Conventional systems handle a rough setting by using various techniques to induce a hypothesis that describes the evidence as well as possible. They aim to maximize the correct cover of the induced hypothesis by maximizing the number of positive examples covered and negative examples not covered. This means that most of the positive evidence would be described, along with some of the negative evidence. The induced hypothesis cannot say with certainty whether an example definitely belongs to the evidence or not. Simultaneously, we also have finished the relative previous research works (Gao *et al.*, 2009, 2010; Wei *et al.*, 2007a,b, 2008a, b, 2009). The objective of the study is concerning the new method implementation of Rough set in ad-hoc networks scenarios.

ROUGH SET AND DISTANCE LEARNING ALGORITHM

Pawlak (1991) set theory from the proposed in the early eighties is a problem of dealing with ambiguity and uncertainty of new mathematical tools. Given a set of objects with different attributes to the information table,

rough set theory allows us to, respectively, in accordance with the conditions of property, making property classification. Each equivalence class based on decision attribute defines a concept.

We use $Des(X_i)$ to state the expression of Equivalence class X_i , namely, it is the set of property values that states attribute set X_i .

Y is a conception, we define the following the lower approximation \underline{Y} and upper approximation \overline{Y} :

$$Y = \{e \in U | e \in X_i, \exists X_i \subset Y\} \quad \overline{Y} = \{e \in U | \exists e \in X_i, \exists X_i \cap Y \neq \emptyset\}$$

If one element belongs to $\overline{Y} - \underline{Y}$ we cannot assure if it exists belong to the domain of Y then we introduce a concept of the discriminant index to measure the extent of this element, which is defined as follows:

$$\alpha_{D_i}(Y) = 1 - \frac{|\overline{Y} - \underline{Y}|}{|U|}$$

For a given information table, not all of the properties are classified by the necessary conditions for the existence of a minimal set of attributes, we get the minimal subset of the process is a reduction process. In this study we applied the relative reduction method (Wei *et al.*, 2010a, b).

ALGORITHM FOR IMPLEMENTATION

Algorithm for distance education, with emphasis on the concept, namely, $Des(Y) = \{\text{Fail}\}$, this is what we want to find out what caused the online students cannot pass the exam. The algorithm can be summarized as the following four steps:

- Step 1:** Calculate the reduction
- Step 2:** Initialize variables, as long as the number of attributes is not zero, then the third step
- Step 3:** When circulating implementation. When circulating in the calculation of condition attributes in the first discriminated index, the next highest index value is stored in the results, and the condition attribute (Hongyan and Maguire, 2003) is removed from the conditions of property concentration. The next output of a decisive or conclusive decision rules to determine a new universe. Basis of this cycle continues, if the new domain of space, then the next step
- Step 4:** Output of all critical and non-conclusive decision rules

Algorithm implementation: Algorithm implementation issues in the discussion, we first introduce the problem

Table 1: Quiz test contrasting

	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Quiz 6	Final
S1	98	100	90	89	91	85	90
S2	100	90	30	45	55	32	40
S3	68	70	80	89	91	85	85
S4	88	100	80	69	75	85	81
S5	76	65	50	70	46	49	46
S6	58	70	90	60	78	97	87

Column means quiz score, S1-S8 means specific course

Table 2: Transfer record contrasting

	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Quiz 6	Final	Total
e1	P	P	P	P	P	P	P	76
e2	P	P	F	F	P	F	F	4
e3	P	F	F	P	P	F	F	3
e4	P	P	F	P	P	P	P	6
e5	P	P	F	P	P	F	F	1
e6	P	P	P	F	F	F	F	5
e7	F	P	P	P	P	P	P	8
e8	P	P	P	P	P	F	P	10

P: Pass, F: Failure

through an example, given a student (Wong *et al.*, 1986) included various quizzes and final exam of the text file. Did not pass the final exam in this course means that did not pass. Therefore, the main issue is, which part of the students lack understanding of the content (i.e., no small achievement tests passed) led to their final exam fail. We can use a table of information to determine (Liang *et al.*, 2000) the final grade is not good cause the relevant rules. In the future so that we can tell the students the teaching of courses in a central part and hence the appropriate guidance to students.

The following is a quiz about the students and the final examination of the text file. Including 115 students, 6 quizzes and a final exam (Table 1).

First, we only care about the students in this course is passed, so we can convert out this form, the table P that pass, F, said failure. Students in the same situation and put into one category which are at a simple table, a total of eight categories, not difficult to find, in the total number of students into a 113, the other two in each quiz are not passed students, we are not considered because of their logical end result should be a Failure (Table 2). In Table 2, P stands for pass, F stands for failure, "e" stands for the number of exam.

Reduction, which requires reduction of redundant information out, in order to achieve this step, we use linear searching and sorting methods to complete. In the table above, for example, the table $R = \{Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, \text{Final}\}$

R positive domain is:

$$POS_R(F) = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}$$

To calculate the reduction on, we need to study whether it is-independent. When we used to test the removal of circulating were the positive field, if every time

Table 3: Simplify data contrasting

	Q3	Q5	Q6	Final	Total
e1	P	P	P	P	76
e2	F	P	F	F	4
e3	F	P	F	F	3
e4	F	P	P	P	6
e5	F	P	F	F	1
e6	P	F	F	F	5
e7	P	P	P	P	8
e8	P	P	P	P	10

P: Pass, F: Failure

a small test field is not equal to the final exam is the positive region, we reserve the condition attributes, the following is a partial result of the operation:

$$\begin{aligned}
 POS_{R-\{Q_1\}}(F) &= \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\} = POS_R(F) \\
 U/IND(R-\{Q_1\}) &= \{\{e_1, e_7\}, \{e_2\}, \{e_3\}, \{e_4\}, \{e_5\}, \{e_6\}, \\
 &\quad \{e_8\}\} \\
 &\dots\dots\dots \\
 U/IND(R-\{Q_6\}) &= \{\{e_1, e_7, e_8\}, \{e_2, e_3, e_4, e_5\}, \{e_6\}\} \\
 POS_{R-\{Q_6\}}(F) &= \{e_1, e_6, e_7, e_8\} \neq POS_R(F)
 \end{aligned}$$

We conclude from the above operation: e_3, e_5, e_6 is essential, because it is not equal to the Final of the domain is the domain. Further, we have simplified information Table 3.

The remaining work is to find decision rules. We have the following results On the field:

$$U = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}$$

- Failure category: $Y = \{e_2, e_3, e_5, e_6\}$
- Based on the non-recognition category Quiz:

$$\begin{aligned}
 Y_1 &= \{e_1, e_6, e_7, e_8\} Des(Y_1) = \{\text{Quiz} = P\} \\
 Y_2 &= \{e_2, e_3, e_4, e_5\} Des(Y_2) = \{\text{Quiz} = F\}
 \end{aligned}$$

- Lower approximation $\underline{Y} = \{\}$
- Upper approximation $\bar{Y} = \{e_1, e_6, e_7, e_8, e_2, e_3, e_4, e_5\}$
- Identify indicators:

$$\alpha_{Q_3} = 1 - \frac{|\bar{Y}/Y|}{|U|} = 0.00$$

Calculated the same Q_5, Q_6 index, we get the right table.

We calculate the α exponentiate and get the right Table 4.

It can be seen Quiz 6 value is the highest recognition that is it is a member of the best decisions in the areas Y, therefore, To the following decision rules we have:

$$\text{Rule 1: } \{\text{Quiz 6} = F\} \Rightarrow \{\text{Final} = F\}$$

The resulting decision tree is as follows:

Table 4: Data contrasting

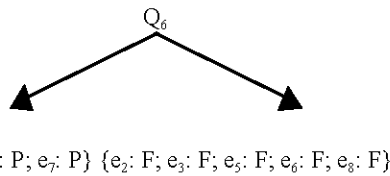
Attribute	α -Value
Q ₃	0.00
Q ₅	0.125
Q ₆	0.375

Column means quiz attribute, α Value means sampling results

Table 5: Record contrasting based on the decision tree

	Q3	Q5	Q6	Final	Total
e2	F	P	F	F	4
e3	F	P	F	F	3
e5	F	P	F	F	1
e6	P	F	F	F	5
e8	P	P	F	P	10

P: Pass, F: Failure



We have been based on decision tree. See also Table 5.

Then we combine with the remaining condition attributes to find the highest recognition of the financial indicator, thus creating two combinations, the right table is the result of the second round of calculations.

Then we combine Q₆ with the remaining condition attributes to find the highest recognition of the financial indicator, thus creating two combinations Q₆, Q₃ and Q₆, Q₅, the right table is the result of the second round of calculations:

Attribute	α
Q ₆ and Q ₃	0.60
Q ₆ and Q ₅	0.20

Not difficult to see Q₆ and Q₃ from the table, with a high recognition index, we get:

The second Rules:

Rule 2: {Quiz 6 = F, Quiz 3 = F} \Rightarrow {Final = F}

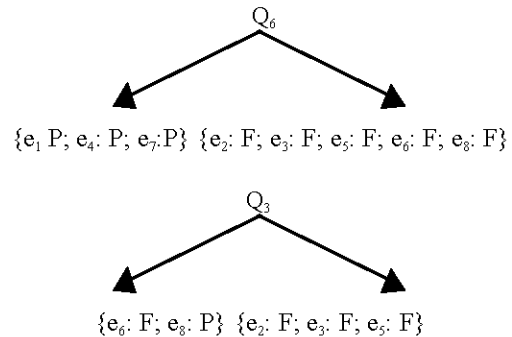
Combination of decision tree as shown on page Table 1.

Using the same method we can get:

Rule 3: {Quiz 6 = F, Quiz 5 = F} \Rightarrow {Final = F}

Over the course of our out of the rules for online students is concerned, these rules have little credibility it? The concept of classification of intensity answered this question, it could use the formula to calculate the:

Rules cover all positive examples
Rules cover all examples(including negative, positive)



In order to calculate the final step, we again used the line search, the rule strength calculation table below (Table 6).

Table 6 can be summarized as follows: If a student does not pass the quiz, then he will have a 56.52% final exam can not be possible. If an online learning and the students did not pass the exam Quiz 6 and Quiz 5, then he can not say 100% pass the final exam. This form of online is learned by the students before the performance analysis, for the later study the same course provides guidance for students.

DECISION-MAKING ROLE IN MEDICAL DIAGNOSIS

Molodtsov (1999) proposed soft set theory to solve the problem of uncertainty, a common method and its non-binding parameters, improving the practical application of the decision problem. It is defined as follows:

Definition 1: U is the initial domain of, E is Parameter set, then (F, E) is U On the soft set if and only if F is from E to U a map of all subsets.

For any $\epsilon \in E$, $e \in U$, $F(\epsilon)$ can be seen as a soft set (F, E) each ϵ has one element. Soft collection of parameter settings without any constraints, the elements can be words, sentences, equations and even functions, mapping. Some of the definitions on rough sets can be defined in parallel over. Consider the soft set (F, E), $P \subset E$, (F, P) is (F, E) a subset of soft set, Now according to the definition of rough sets to define reduction of soft set, respectively and the nuclear soft set. If Q is P reduction, then soft set (F, Q) is soft set (F, P) the reduction of soft sets, if C is P's shore, then soft set (F, C) is (F, P) nuclear soft set.

Maji *et al.* (2002) have used a soft set to make home buying decision problem, also in the clinical diagnosis, we can also do the decision-making problems with soft set. For Table 1, the initial universe is to be considered in

