Volume 1, No. 2, July-August 2010

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

A Novel Approach of Fuzzy Set Analysis in Distributed Data Mining

Vuda Sreenivasa Rao*
Research Scholar, CSIT Department
JNT University, Hyderabad
Andhra Pradesh, India.
vudasrinivasarao@gmail.com

Dr. S Vidyavathi
Associate Professor, CSIT Department
JNT University, Hyderabad
Andhra Pradesh, India.
vidyasom@yahoo.co.in

Abstract: Distributed Data Mining(DDM) has evolved into an important and active area of research because of theoretical challenges and practical applications associated with the problem of extracting, interesting and previously unknown knowledge from very large real-world databases. Fuzzy Set Theory (FST) is a mathematical formalism for representing uncertainty that can be considered an extension of the classical set theory. It has been used in many different research areas, including those related to inductive machine learning and reduction of knowledge in Distributed data-based systems. One important concept related to FST is that of a fuzzy relation. In this paper we presented the current status of research on applying fuzzy set theory to DDM, which will be helpful for handle the characteristics of real-world databases. The main aim is to show how fuzzy set and fuzzy set analysis can be effectively used to extract knowledge from large databases.

Keywords: Data mining, Data tables, Distributed Data Mining (DDM), Fuzzy sets, Fuzzy set analysis.

I. INTRODUCTION

Data mining technology has emerged as a means for identifying patterns and trends from large quantities of data. Data mining is a computational intelligence discipline that contributes tools for data analysis, discovery of new knowledge, and autonomous decision making. The task of processing large volume of data has accelerated the interest in this field. As mentioned in Mosley (2005) data mining is the analysis of observational datasets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner.

Distributed Data Mining (DDM) aims at extraction useful pattern from distributed heterogeneous data bases in order, for example, to compose them within a distributed knowledge base and use for the purposes of decision making. A lot of modern applications fall into the category of systems that need DDM supporting distributed decision making. Applications can be of different natures and from different scopes, for example, data and information fusion for situational awareness; scientific data mining in order to compose the results of diverse experiments and design a model of a phenomena, intrusion detection, analysis, prognosis and handling of natural and man-caused disaster to prevent their catastrophic development, Web mining ,etc. From practical point of view, DDM is of great concern and ultimate urgency.

Fuzzy set theory has been studied extensively over the past 30 years. Most of the early interest in fuzzy set theory pertained to representing uncertainty in human cognitive processes (see for example Zadeh (1965)). Fuzzy set theory is now applied to problems in engineering, business, medical and related health sciences, and the natural sciences.

The use of fuzzy set theory as a methodology for modeling and analyzing decision systems is of particular interest to researchers in production management due to fuzzy set theory's ability to quantitatively and qualitatively model problems which involve vagueness and imprecision. Kawasaki and Evans (1986) identify the potential applications of fuzzy set theory to the following areas of production management: new product development, facilities location and layout, production scheduling and control, inventory management, quality and cost benefit analysis. Kawasaki and Evans identify three key reasons why fuzzy set theory is relevant to production management research. First, imprecision and vagueness are inherent to the decision maker's mental model of the problem under study. Thus, the decision maker's experience and judgment may be used to complement established theories to foster a better understanding of the problem. Second, in the production management environment, the information required to formulate a model's objective, decision variables, constraints and parameters may be vague or not precisely measurable. Third, imprecision and vagueness as a result of personal bias and subjective opinion may further dampen the quality and quantity of available information.

Fuzzy set theory is a new mathematical approach to imperfect knowledge. The problem of imperfect knowledge has been tackled for a long time by philosophers, logicians and mathematicians. Recently it became also a crucial issue for computer scientists, particularly in the area of artificial intelligence. There are many approaches to the problem of how to understand and manipulate imperfect knowledge. The most successful one is, no doubt, the fuzzy set theory proposed by Zadeh .Fuzzy set theory proposed by Zpawlak presents still another attempt to this problem. The theory has attracted attention of many researchers and practitioners all over the world, who contributed essentially to its development and applications. Fuzzy set theory has an overlap with many other theories. However we will refrain to discuss these connections here. Despite of the above mentioned connections fuzzy set theory may be considered as the independent discipline in its own rights. Fuzzy set theory has found many interesting applications. The fuzzy set approach seems to be of fundamental importance to AI and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, inductive reasoning and pattern recognition. The main advantage of fuzzy set theory in data analysis is that it does not need any preliminary or additional information about data – like probability in statistics, or basic probability assignment in Dempster-Shafer theory, grade of membership or the value of possibility in rough set theory.

The proposed approach

A. Provides efficient algorithms for finding hidden patterns in data,

B. Finds minimal sets of data (data reduction),

C. Evaluates significance of data,

D. Generates sets of decision rules from data,

E. It is easy to understand,

F. Offers straightforward interpretation of obtained results,

G. Most algorithms based on the rough set theory are particularly suited for parallel processing.

The remaining sections of the paper are organized as follows. In Section II we describe Fuzzy sets theory in data mining .In Section III we describe Distributed Data Mining.

In Section IV we describe fuzzy set analysis in Distributed Data Mining. In Section V we describe Computational Aspects of Fuzzy set on DDM. In Section VI concludes the paper.

II.FUZZY SETS THEORY IN DATA MINING

Fuzzy set theory was developed by Zdzislaw Pawlak in the early 1980's1. Fuzzy set deals with classification of discreet data table in a supervised learning environment. Although in theory fuzzy set deals with discreet data, fuzzy set is commonly used in conjunction with other technique to do discrimination on the dataset. The main feature of fuzzy set data analysis is non-invasive, and the ability to handle qualitative data. This fits into most real life application nicely. Fuzzy set have seen light in many researches but seldom found its way into real world application.

Knowledge discovery with fuzzy set is a multi-phase process consisted of mainly:

A. Discretization

B. Reducts and rules generation on training set

A. System architecture

The overview of the architecture of the system can be seen in figure. The proposed architecture will adopt the traditional architecture of a data mining system. Data from multiple channels is collected on the operational data store for fast transaction and up to date data that can be used for the front office. Then, periodically, the data is extracted, cleans, transformed and imported into the data warehouse. The data will then will be send to the appropriate data marts for departmental use. Then, according to the needs of the user, either the enterprise data or the departmental data is sent to the OLAP tier for processing. The results is then stored and then sent to the decision makers through the use of thin clients. The overview of this architecture is seen in Figure 1 . The proposed system is pretty good in theory as it provides compartmentalization of data and collection of data from multiple channels. The architecture is simple and sticks to the basis of founded work and should provide a good base for the system.

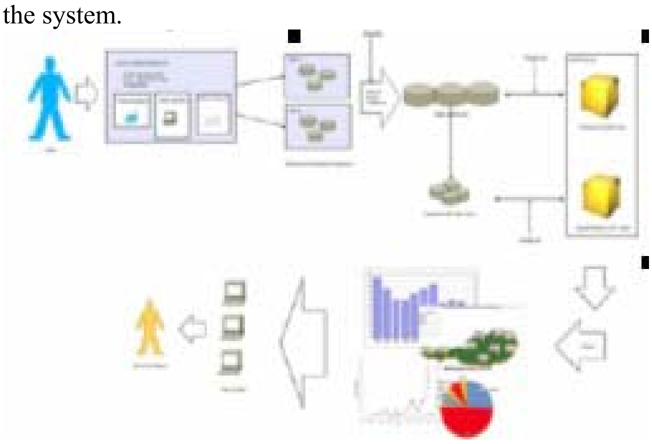


Figure 1: the over view of proposed system Architecture

III. DISTRIBUTED DATA MINING

Traditional warehouse-based architectures for data mining suppose to have centralized data repository. Such a centralized approach is fundamentally inappropriate for most of the distributed and ubiquitous data mining applications. In fact, the long response time, lack of proper use of distributed resource, and the Fundamental characteristic of centralized data mining algorithms do not work well in distributed environments. A scalable solution for distributed applications calls for distributed processing of data, controlled by the available resources and human factors. For example, let us suppose an ad hoc wireless sensor network where the different sensor nodes are monitoring some time-critical events. Central collection of data from every sensor node may create traffic over the limited bandwidth wireless channels and this may also drain a lot of power from the devices. A distributed architecture for data mining is likely aimed to reduce the communication load and also to reduce the battery power more evenly across the different nodes in the sensor network. One can easily imagine similar needs for distributed computation of data mining primitives in ad hoc wireless networks of mobile devices like PDAs, cell phones, and wearable computers. The wireless domain is not the only example. In fact, most of the applications that deal with time-critical distributed data are likely to benefit by paying careful attention to the distributed resources for computation, storage, and the cost of communication. As an other example, let us consider the World Wide Web: it contains distributed data and computing resources. An increasing number of databases (e.g., weather databases, oceanographic data, etc.) and data streams (e.g., financial data, emerging disease information, etc.) are currently made on-line, and many of them change frequently. It is easy to think of many applications that require regular r monitoring of these diverse and distributed sources of data. A distributed approach to analyze this data is likely to be more scalable and practical particularly when the application involves a large number of data sites. Hence, in this case we need data mining architectures that pay careful attention to the distribution of data, computing and communication, in order to access and use them in a near optimal fashion. Distributed Data Mining (sometimes referred by the acronym DDM) considers data mining in this broader context. DDM may also be useful in environments with multiple compute nodes connected over high speed networks. Even if the data can be quickly centralized using the relatively fast network, proper balancing of computational load among a cluster of nodes may require a distributed approach. The privacy issue is playing an increasingly important role in the emerging data mining applications. For example, let us suppose a consortium of different banks collaborating for detecting frauds. If a centralized solution was adopted, all the data from every bank should be collected in a single location, to be processed by a data mining system. Nevertheless, in such a case a Distributed Data Mining system should be the natural technological choice: both it is able to learn models from distributed data without exchanging the raw data between different repository, and it allows detection of fraud by preserving the privacy of every bank's customer transaction data. For what concerns techniques and architecture, it is worth noticing that many several other fields influence Distributed Data Mining systems concepts. First, many DDM systems adopt the Multi-Agent System (MAS) architecture, which finds its root in the Distributed Artificial Intelligence (DAI). Second, although Parallel Data Mining often assumes the presence of high speed met work connections among the computing nodes, the development of DDM has also been influenced by the PDM literature. Most DDM algorithms are designed upon the potential parallelism they can apply over the given distributed data.. In figure 3 a general Distributed Data Mining framework is presented. In essence, the success of DDM algorithms lies in the aggregation. Each local model represents locally coherent patterns, but lacks details that may be required to induce globally meaningful knowledge. For this reason, many DDM algorithms require a centralization of a subset of local data to compensate it. The ensemble approach has been applied in various domains to increase the accuracy of the predictive model to be learnt. It produces multiple models and combines them to enhance accuracy. Typically, voting (weighted or un-weighted) schema are employed to aggregate base model for obtaining a global model. As we have discussed above, minimum data transfer is another key attribute of the successful DDM algorithm.

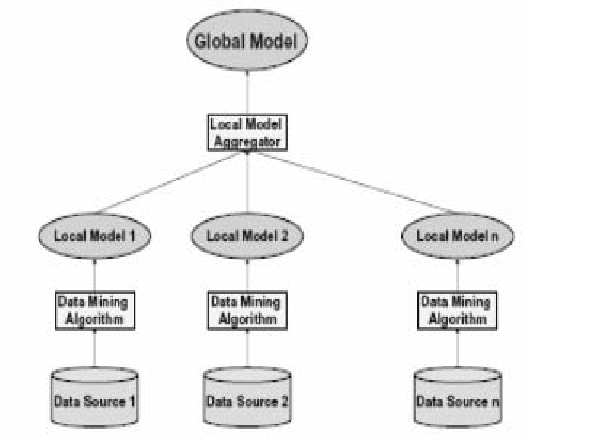


Figure 2: General Distributed data mining Frame work

IV.FUZZY SET ANALYSIS IN DISTRIBUTED DATA MINING

A. Fuzzy sets

Fuzzy sets have been introduced by Lotfi A. Zadeh (1965). What Zadeh proposed is very much a paradigm shift that first gained acceptance in the Far East and its successful application has ensured its adoption around the world. Fuzzy sets are an extension of classical set theory and are used in fuzzy logic. In classical set theory the membership of elements in relation to a set is assessed in binary terms according to a crisp condition — an element either belongs or does not belong to the set. By contrast, fuzzy set theory permits the gradual assessment of the membership of elements in relation to a set; this is described with the aid of a membership function valued in the real unit interval [0, 1]. Fuzzy sets are an extension of classical set theory since, for a certain universe, a membership function may act as an indicator function, mapping all elements to either 1 or 0, as in the classical notion.

Specifically, A fuzzy set is any set that allows its members to have different grades of membership (membership function) in the interval [0,1]. A fuzzy set on a classical set X is defined as follows:

$$\tilde{A} = \{(x, \mu_A(x)) \mid x \in X\}$$

B. The membership functions in fuzzy set analysis

The membership function $\mu_A(x)$ quantifies the grade of membership of the elements x to the fundamental set X. An element mapping to the value 0 means that the member is not included in the given set, 1 describes a fully included member. Values strictly between 0 and 1 characterize the fuzzy members.

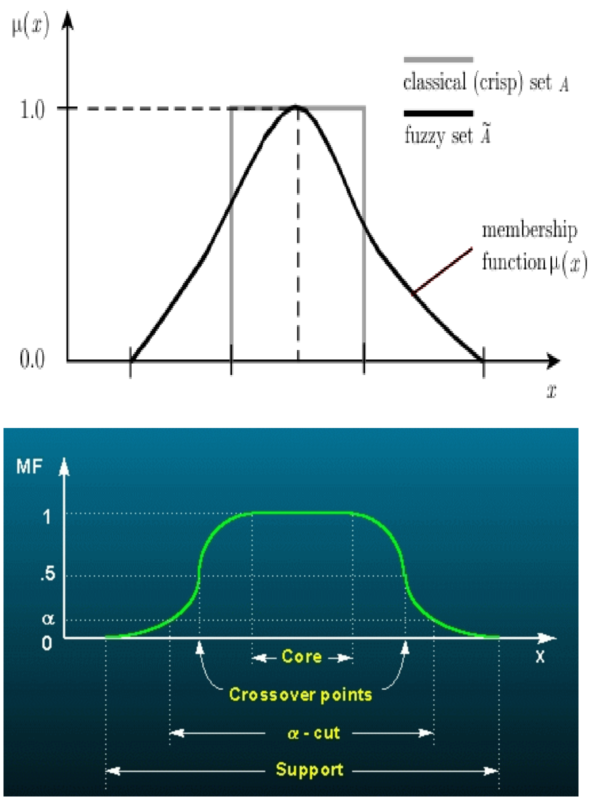


Figure 3

C. Membership function terminology:

Universe of Discourse: the universe of discourse is the range of all possible values for an input to a fuzzy system.

Support: the support of a fuzzy set F is the crisp set of all points in the universe of discourse U such that the membership function of F is non-zero.

$$\text{Supp}A = \{x \mid \mu_A(x) > 0,\ \forall x \in \mathrm{X}\}$$

Core: the core of a fuzzy set F is the crisp set of all points in the universe of discourse U such that the membership function of F is 1.

$$\text{core}A = \{x \mid \mu_A(x) = 1,\ \forall x \in \mathrm{X}\}$$

Boundaries: the boundaries of a fuzzy set F is the crisp set of all points in the universe of discourse U such that the membership function of F is between 0 and 1.

$$\text{Boundaries}A = \{x \mid 0 < \mu_A(x) < 1,\ \forall x \in \mathrm{X}\}$$

Crossover point: the crossover point of a fuzzy set is the element in U at which its membership function is 0.5.

$$\mu(x) = 0.5$$

Height: the biggest value of membership functions of fuzzy set.

Normalized fuzzy set: the fuzzy set of $\text{Height}(A) = 1$

Cardinality of the set: X : *finite*

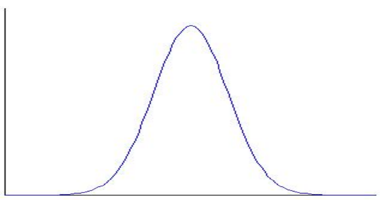
$$|A| = \sum_{x\in \mathrm{X}} \mu_{\mathrm{A}}(x) = \sum_{x\in \mathrm{Supp}(A)} \mu_{\mathrm{A}}(x)$$

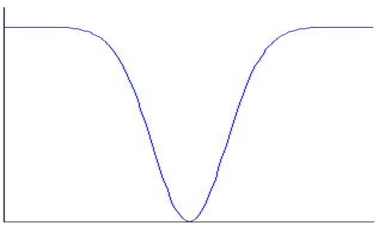
Relative cardinality: $\|A\| = \dfrac{|A|}{|\mathrm{X}|}$

Convex fuzzy set: $\mathrm{X} \in \mathrm{R}$,

a fuzzy set A *is* Convex, if for $\forall \lambda \in [0,\ 1]$

$$\mu_A(\lambda x_1 + (1-\lambda)x_2) \ge \min(\mu_A(x_1),\ \mu_A(x_2))$$





D. Type of membership functions:

1. Numerical definition (discrete membership functions)

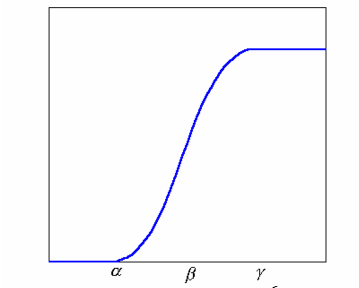
$$A = \sum_{x_i \in X} \mu_A(x_i)/x_i$$

2. Function definition (continuous membership functions)

Including of S function, Z Function, Pi function, Triangular shape, Trapezoid shape, Bell shape.

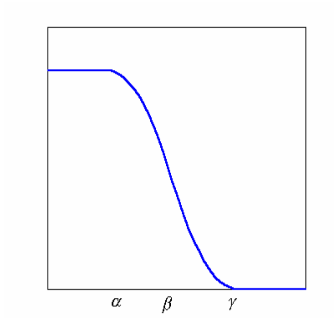
$$A = \int_X \mu_A(x)/x$$

(1) **S function**: monotonically increasing membership function



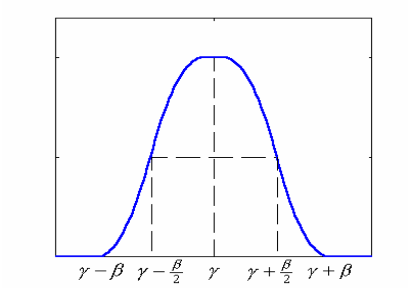
$$S(x;\alpha,\beta,\gamma) = \begin{cases} 0 & for\ x \le \alpha \\ 2(\frac{x-\alpha}{\gamma-\alpha})^2 & for\ \alpha \le x \le \beta \\ 1-2(\frac{x-\alpha}{\gamma-\alpha})^2 & for\ \beta \le x \le \gamma \\ 1 & for\ \gamma \le x \end{cases}$$

(2) **Z function:** monotonically decreasing membership function



$$Z(x;\alpha,\beta,\gamma) = \begin{cases} 1 & for\ x \le \alpha \\ 1-2(\frac{x-\alpha}{\gamma-\alpha})^2 & for\ \alpha \le x \le \beta \\ 2(\frac{x-\alpha}{\gamma-\alpha})^2 & for\ \beta \le x \le \gamma \\ 0 & for\ \gamma \le x \end{cases}$$

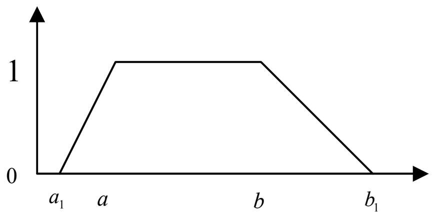
(3) **Π Function:** combine S function and Z function, monotonically increasing and decreasing membership function



$$\Pi(x;\beta,\gamma) = \begin{cases} S(x;\ \gamma-\beta,\ \gamma-\frac{\beta}{2},\ \gamma) & for\ x \le \gamma \\ 1-S(x;\ \gamma,\ \gamma+\frac{\beta}{2},\ \gamma+\beta) & for\ x \ge \gamma \end{cases}$$

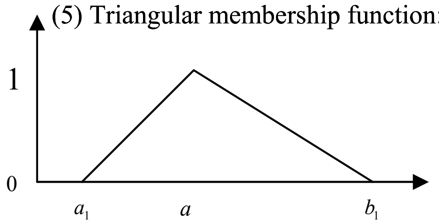
Piecewise continuous membership function

(4)Trapezoidal membership function:



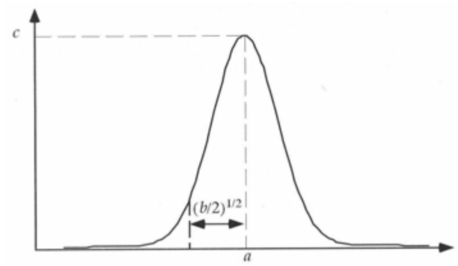
$$\mu_A(x) = \begin{cases} 0 & for \ x \le a_1 \\ \frac{x-a_1}{a-a_1} & for \ a_1 \le x \le a \\ 1 & for \ a \le x \le b \\ \frac{b_1-x}{b_1-b} & for \ b \le x \le b_1 \\ 0 & for \ b_1 \le x \end{cases}$$

(5) Triangular membership function:



$$\mu_A(x) = \begin{cases} 0 & for \ x \le a_1 \\ \frac{x-a_1}{a-a_1} & for \ a_1 \le x \le a \\ \frac{b_1-x}{b_1-a} & for \ a \le x \le b_1 \\ 0 & for \ b_1 \le x \end{cases}$$

(6) Bell-shaped membership function:



Before illustrating the mechanisms which make fuzzy logic machines work, it is important to realize what fuzzy logic actually is. Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth- truth values between "completely true" and "completely false". As its name suggests, it is the logic underlying modes of reasoning which are approximate rather than exact. The importance of fuzzy logic derives from the fact that most modes of human reasoning and especially common sense reasoning are approximate in nature.
The essential characteristics of fuzzy logic are as follows.

A. In fuzzy logic, exact reasoning is viewed as a limiting case of approximate reasoning.
   B. In fuzzy logic everything is a matter of degree.
   C. Any logical system can be fuzzified.
   D. In fuzzy logic, knowledge is interpreted as a collection of elastic or, equivalently, fuzzy constraint on a collection of variables.
   E. Inference is viewed as a process of propagation of elastic constraints.

After know about the characteristic of fuzzy set, we will introduce the operations of fuzzy set. A fuzzy number is a convex, normalized fuzzy set $\tilde{A} \subseteq \mathbb{R}$ whose membership function is at least segmental continuous and has the functional value $\mu_A(x) = 1$ at precisely one element. This can be likened to the funfair game "guess your weight," where someone guesses the contestants weight, with closer guesses being more correct, and where the guesser "wins" if they guess near enough to the contestant's weight, with the actual weight being completely correct (mapping to 1 by the membership function). A fuzzy interval is an uncertain set $\tilde{A} \subseteq \mathbb{R}$ with a mean interval whose elements possess the membership function value $\mu_A(x) = 1$. As in fuzzy numbers, the membership function must be convex, normalized, and at least segmental continuous.

Set- theoretic operations

Subset: $A \subseteq B \Leftrightarrow \mu_A \le \mu_B$

Complement: $\overline{A} = X - A \Leftrightarrow \mu_{\overline{A}}(x) = 1 - \mu_A(x)$

Union:

$$C = A \cup B \Leftrightarrow \mu_c(x) = \max(\mu_A(x), \mu_B(x)) = \mu_A(x) \vee \mu_B(x)$$

Intersection:

$$C = A \cap B \Leftrightarrow \mu_c(x) = \min(\mu_A(x), \mu_B(x)) = \mu_A(x) \wedge \mu_B(x)$$

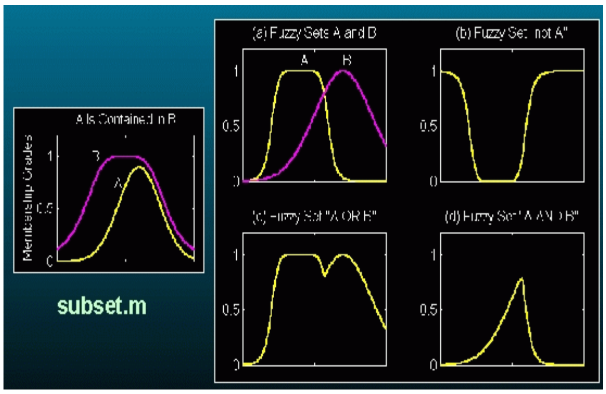


Figure 4

Although one can create fuzzy sets and perform various operations on them, in general they are mainly used when creating fuzzy values and to define the linguistic terms of fuzzy variables. This is described in the section on fuzzy variables. At some point it may be an interesting exercise to add fuzzy numbers to the toolkit. These would be specializations of fuzzy sets with a set of operations such as addition, subtraction, multiplication and division defined on them.

According to the characteristics of triangular fuzzy numbers and the extension principle put forward by Zadeh (1965), the operational laws of triangular fuzzy numbers, $\tilde{A} = (l_1, m_1, r_1)$ and $\tilde{B} = (l_2, m_2, r_2)$ are as follows:

(A) Addition of two fuzzy numbers

$$(l_1, m_1, r_1) \oplus (l_2, m_2, r_2) = (l_1 + l_2, m_1 + m_2, r_1 + r_2)$$

(B) Subtraction of two fuzzy numbers

$$(l_1, m_1, r_1) \Theta (l_2, m_2, r_2) = (l_1 - r_2, m_1 - m_2, r_1 - l_2)$$

(C) Multiplication of two fuzzy numbers

$$(l_1, m_1, r_1) \otimes (l_2, m_2, r_2) \cong (l_1 l_2, m_1 m_2, r_1 r_2)$$

(D) Division of two fuzzy numbers

$$(l_1, m_1, r_1) \varnothing (l_2, m_2, r_2) \cong (l_1 / r_2, m_1 / m_2, r_1 / l_2)$$

When we through the operations of fuzzy set to get the fuzzy interval, next we will convert the fuzzy value into the crisp value. Below are some methods that convert a fuzzy set back into a single crisp (non-fuzzy) value. This is something that is normally done after a fuzzy decision has been made and the fuzzy result must be used in the real world. For example, if the final fuzzy decision were to adjust the temperature setting on the thermostat a 'little higher', then it would be necessary to convert this 'little higher' fuzzy value to the 'best' crisp value to actually move the thermostat setting by some real amount.

V. COMPUTATIONAL ASPECTS OF FUZZY SET ON DDM

In the literature, there has long been a lack of time complexity analysis of algorithms for frequently used fuzzy set operations. Time complexities of constructing an equivalence relation are shown to be $O(lm2)$, where l and m are number of attributes and objects, respectively . This result corresponds to the analysis of an algorithm, reported in, where the goal is to obtain the equivalence relation according to the values of a single attribute. For a given functional dependency $X \rightarrow Y$ that holds in an information table S, we say that $x \rightarrow X$ is superfluous (or non-significant) attribute for Y in S if and only if, X-$\{x\} \rightarrow Y$ still holds in S. A reduct of X for Y in S is a subset P of X such that P does not contain any superfluous attribute. If we have a metric to measure the degree of dependency, then we have a way to explore a reduct of X, with a degree of θ, where $0 \leq \theta \leq 1$. It is shown that finding a reduct of X for Y in S is computationally bounded by $l2m2$ where l and m is a length of X and the number of objects in S respectively. The time complexity to find all reducts of X is $O(2lJ)$, where J is the computational cost for finding one reduct, and l is the number of attributes in X.

VI. CONCLUSIONS

In the paper, basic concepts of distributed data mining and the fuzzy set theory were discussed. Fuzzy Set Theory has been widely used in DDM since it was put forward. Having important functions in the expression, study, conclusion and etc. of the uncertain knowledge, it is a powerful tool, which sets up the intelligent decision system. The main focus is to show how fuzzy set techniques can be employed as an approach to the problem of data mining and knowledge extraction. The project shows that fuzzy set theory can be used as a tool for knowledge discovery. Even though it is a symbolical method, application of a suitable quantization technique will allow it to perform on just about any type of data. As opposed to numerical method that cannot be adapted to be used for symbolical data. Fuzzy set provide a useful tool that can be used on a lot of different data regardless weather it is numerical or symbolical and it also provide a non-intrusive methodology to knowledge discovery.

VII. REFERENCES

[1] Amin, K. I. (1993). Jump diffusion option valuation in discrete time. Journal of Finance, 48(5), 1833–1863.

[2] Barone-Adesi, G. and R. E. Whaley (1987). Efficient analytic approximation of American option values. Journal of Finance, 42(2), 301–320.

[3] Black, F., and M. Scholes (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637–654.

[4] Brennan, M. J. and E. S. Schwartz (1977). The valuation of American put options. Journal of Finance, 32(2), 449–462.

[5] Cox, J. C. and S. A. Ross (1975). Notes on option pricing I: Constant elasticity of variance diffusion. Working paper. Stanford University.

[6] Cox, J. C., S. A. Ross, and M. Rubinstein (1979). Option pricing: A simplified approach. Journal of Financial Economics, 7(3), 229–263.

[7] Lee, C.F., G.H. Tzeng, and S.Y. Wang (2005). A new application of fuzzy set theory to the Black-Scholes option pricing model. Expert Systems with Applications, 29(2), 330-342.

[8] Lee, C.F., G.H. Tzeng, and S.Y. Wang (2005). A Fuzzy set approach to generalize CRR model: An empirical analysis of S&P 500 index option. Review of Quantitative Finance and Accounting, 25(3), 255-275.

[9] Lee, J. C., C. F. Lee, and K. C. J. Wei (1991). Binomial option pricing with stochastic parameters: A beta distribution approach. Review of Quantitative Finance and Accounting, 1(3), 435–448.

[10] Goguen, J. A. (1967). *L*-fuzzy sets. Journal of Mathematical Analysis and Applications, 18, 145–174.

[11] Geske, R. and H. E. Johnson (1984). The american put valued analytically. Journal of Finance, 1511–1524.

[12] Gottwald, S. (2001). A Treatise on Many-Valued Logics. Baldock, Hertfordshire, England: Research Studies Press Ltd.

[13] Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. Journal of Finance, 42(2), 281–300.

[14] Klir, G.J. and B. Yuan. (1995). Fuzzy Sets and Fuzzy Logic. Theory. and Applications, Ed. Prentice-Hall.

[15] Scott, L. (1987). Option pricing when variance changes randomly: Theory, estimation and an application. Journal of Financial and Quantitative Analysis, 22(4), 419–438.

[16] Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical evidence. Journal of Financial Economics, 19(2), 351–372.

[17] Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8, 338–353.

[18] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning. Information Sciences, 8,199–249, 301–357; 9, 43–80.

[19] Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3–28.

[20] vuda sreenivasarao (2010) A novel approach of rough analysis in distributed data mining, volm4,1237-1244.

Authors:

Vuda Sreenivasarao received the M.Tech degree in Computer Science & Engg from the Satyabama University, in 2007. . He is research scholar in CSIT Department, JNT University Hyderabad Andhra Pradesh, India. His research interests include Network Security, Cryptography, and Data Mining & Artificial Intelligence.

Dr. S Vidyavathi received her PhD degree from IIT Mumbai, She is currently working as Associate Professor in CSIT Department , JNT University, Andhra Pradesh, India.