Fuzzy Versus Quantitative Association Rules: A Fair Data-Driven Comparison

Hannes Verlinde, Martine De Cock, and Raymond Boute

***Abstract*—As opposed to quantitative association rule mining, fuzzy association rule mining is said to prevent the overestimation of boundary cases, as can be shown by small examples. Rule mining, however, becomes interesting in large databases, where the problem of boundary cases is less apparent and can be further suppressed by using sensible partitioning methods. A data-driven approach is used to investigate if there is a significant difference between quantitative and fuzzy association rules in large databases. The influence of the choice of a particular triangular norm in this respect is also examined.**

***Index Terms*—Data mining, fuzzy association rules, quantitative association rules, triangular norms.**

I. Introduction

The discovery of knowledge in databases, also called data mining, is a most promising and important research area. In data mining, association rules are often used to represent and identify dependencies between attributes in a database. The original idea dates back to the late 1970s [1], while its application for market basket analysis gained popularity at the beginning of the 1990s [2]. In the original context of association rule mining, data are represented by a table with binary values. The rows correspond to objects or transactions, while the columns correspond to attributes or items. The binary values denote whether a transaction contains a specific attribute. The purpose of association rule mining is to detect rules of the form $A \to B$, indicating that a transaction containing attribute A is likely to contain attribute B as well.

In most real-life applications, databases contain many other values besides 0 and 1. Very common, for instance, are quantitative attributes such as age or income, taking values from an ordered numerical scale, often a subset of the real numbers. One way of dealing with a quantitative attribute is to replace it by a few other attributes that form a crisp partition of the range of the original one. For instance, in a particular application, we might decide to replace age by the attributes young, middle-aged, and old corresponding to intervals [0,35[, [35,65[, and [65,100], respectively, while income can be replaced by low, medium, and high corresponding to the intervals [0,1000[, [1000,2000[, and [2000,10 000], respectively. The new attributes can be considered as binary ones (e.g., the value of young is 1 if the corresponding value of age belongs to [0,35[; otherwise it is 0), which reduces the problem to traditional association rule mining with binary values. The generated rules are now called quantitative association rules [3].

The starting point for fuzzy set theory [4] is that it is against intuition to model vague concepts such as young and high by crisp intervals. For why would a person be considered as young while he is younger than 35, and on his 35th birthday suddenly lose this status? In reality, the transition between being young and not being young is not abrupt but gradual. This is a very good argument for modeling vague concepts by fuzzy sets instead of crisp sets. Many researchers have already used this argument for the introduction of fuzzy association rules (see, e.g., [5]–[11]). In this process, a database containing quantitative attributes is replaced by one with values from [0,1] in a similar way as one does for quantitative association rules. It is said that, compared to quantitative association rules, fuzzy association rules correspond better to intuition and prevent overestimation of boundary cases (see, e.g., [7], [8], and [11]). The so-called sharp boundary problem states that binary algorithms either ignore or overemphasize the elements near the boundary of the intervals in the mining process. It is easy to construct toy examples to illustrate this point.

Association rule mining is, however, not developed to play around with small toy examples but to deal with large databases. Unfortunately, comparisons for large data sets of results obtained with a quantitative versus a fuzzy association rule mining algorithm are extremely hard to find in the literature. An additional problem is the partitioning of the range of attribute values in intervals (for quantitative association rule mining) and the construction of membership functions (for fuzzy association rule mining). The two extreme solutions to this problem are the expert-driven approach (an expert manually sets the interval boundaries and/or defines the membership functions) and the data-driven approach (they are generated automatically from the data table, see, e.g., [3] and [12]). In [11], results obtained with quantitative and fuzzy association rule mining are compared for two artificially created data sets. For the quantitative case, the attribute partitioning method proposed by Srikant and Agrawal [3] is applied. For the fuzzy case, however, the membership functions are constructed manually (but not given in the paper). This is an unfair footing for comparison because additional expert knowledge is injected into the fuzzy approach while the quantitative approach is fully automatical (data driven). It is clear that a fair comparison can only be done using either the same expert-driven approach or the same data-driven approach for both mining processes. In this paper, we experimentally investigate the latter.

II. Association Rule Mining

Recall that a fuzzy set A on a universe X is characterized by $X - [0, 1]$ mapping, also called the membership function of A. For x in X, $A(x)$ denotes the membership degree of x in the fuzzy set A. If the membership function only takes values in {0,1}, it coincides with the traditional set concept, which, in this context, is also called "crisp set."

Table I presents what could happen if we replace the quantitative attributes in a small database by either binary or fuzzy attributes.

Let X be the set of all transactions (in our example $|X| = 6$). We study rules of the form $A \to B$, where A and B are different attributes.[1] We use $A(x)$ to denote the value of attribute A for transaction x. In this way, A becomes either a crisp subset of X [$A(x) = 0$ corresponds to $x \notin A$ while $A(x) = 1$ means $x \in A$] or a fuzzy set in X [$A(x)$ denotes the membership degree of x in A]. The support and confidence of a (candidate) association rule $A \to B$ are defined as

$$\text{supp}(A \to B) = \frac{|A \cap B|}{|X|}$$

$$\text{conf}(A \to B) = \frac{|A \cap B|}{|A|}.$$

[1]For reasons explained in Section III-C, we only consider rules with one attribute in antecedent and consequent.

Manuscript received September 8, 2004; revised March 11, 2005. This work was supported by the bilateral China-Flanders cooperation project 174BO201. H. Verlinde is a Research Assistant of the Fund for Scientific Research–Flanders. M. De Cock is a Postdoctoral Fellow of the Fund for Scientific Research–Flanders. This paper was recommended by Associate Editor T. Sudkamp.

The authors are with Ghent University, B900 Ghent, Belgium.

Digital Object Identifier 10.1109/TSMCB.2005.860134

TABLE I
DATABASE TOY EXAMPLE. (a) QUANTITATIVE ATTRIBUTES. (b) BINARY ATTRIBUTES. (c) FUZZY ATTRIBUTES

(a)

	Age	Income
x_1	35	900
x_2	36	980
x_3	35	950
x_4	40	1500
x_5	52	1700
x_6	64	3000

(b)

	young	middle–aged	old	low	medium	high
x_1	0	1	0	1	0	0
x_2	0	1	0	1	0	0
x_3	0	1	0	1	0	0
x_4	0	1	0	0	1	0
x_5	0	1	0	0	1	0
x_6	0	1	0	0	0	1

(c)

	young	middle–aged	old	low	medium	high
x_1	0.5	0.5	0	0.8	0.2	0
x_2	0.4	0.6	0	0.7	0.3	0
x_3	0.5	0.5	0	0.7	0.3	0
x_4	0	1	0	0	1	0
x_5	0.1	0.7	0.2	0	0.9	0.1
x_6	0	0.5	0.5	0	0	1

The support assesses the statistical significance of the association rule while the confidence measures its strength. We use the definitions of (Zadeh) intersection and cardinality of fuzzy sets A and B in X, i.e.,

$$(A \cap B)(x) = \min\left(A(x), B(x)\right), \qquad \text{for } x \text{ in } X$$

$$|A| = \sum_{x \in X} A(x).$$

Other common definitions of intersection of fuzzy sets will be investigated in Section III-E.

TABLE II
SUPPORT AND CONFIDENCE

CANDIDATE RULE	BINARY		FUZZY	
	supp	conf	supp	conf
middle–aged→low	0.50	0.50	0.27	0.42
middle–aged→medium	0.33	0.33	0.42	0.66
middle–aged→high	0.17	0.17	0.1	0.16

Table II illustrates that mining association rules starting from binary or fuzzy attributes can make a big difference. For instance, in the "fuzzy version," middle-aged → medium clearly surfaces as a rule with high confidence, while in the "binary version" it is considered to be less confident than middle-aged → low, and it might not even meet the threshold. The difference is mainly caused by the presence of boundary cases in Table I(a) (ages like 35 and 64 that are very close to the extremes of the interval for middle-aged).

This toy example is, however, not representative for the kind of applications where association rule mining is typically used. In real-life applications, real-life data are being used and the influence of boundary cases is less apparent in large databases. Furthermore, in practice, the unwanted effect of boundary cases can be further suppressed by using appropriate partitioning methods that look for intervals with densely populated centers. Hence, the arguments typically made in favor of fuzzy association rule mining do not seem to hold for real-life applications, and we do not expect fuzzy association rule mining to be all that different from plain quantitative association rule mining. We will test this conjecture empirically by doing the job for some large databases that are typically used in association rule mining.

III. EXPERIMENTAL APPROACH

A. Data Sets

In the first series of experiments, we use the database FAM95, containing the results of the March 1995 U.S. Current Population Survey, conducted by the Bureau of Census for the Bureau of Labor Statistics.[2] The database consists of 63 756 family records with 23 attributes from which we extract the following six quantitative attributes: number of persons in the family, number of children, family income, age of head of the family, educational level of head, and head's personal income.

Another data set used in the experiments (which we will call HEMAT) contains the results of 42 915 hematological analyses from 459 patients of hepatitis B and C, who were admitted to Chiba University Hospital in Japan.[3] Again we extract six quantitative attributes: white blood cells count, hemoglobin count, hematocrit count, mean corpuscular volume, mean corpuscular hemoglobin, and mean corpuscular hemoglobin concentration.

In addition to these large databases, we also use the data set STULONG, concerning a study of the risk factors of atherosclerosis

[2]This data set was obtained from the UCLA Statistics Data Sets Archive website http://www.stat.ucla.edu/data/fpp.

[3]This data set was obtained from the ECML/PKDD 2004 Discovery Challenge website http://lisp.vse.cz/challenge/ecmlpkdd2004.

in a population of 1417 middle-aged men.[4] Here, we extract five quantitative attributes out of a total of 64. The selected attributes are height, weight, systolic blood pressure, diastolic blood pressure, and cholesterol level.

B. Data-Driven Partitioning

In order to compare quantitative and fuzzy association rule mining, we need comparable techniques for both a crisp and a fuzzy partitioning of the range of each quantitative attribute. To this end, we use clustering techniques as suggested in [3].

The classical k-means algorithm [13] is an exclusive or crisp clustering algorithm that starts with an initial estimation of the cluster centers and minimizes the sum of squared distances from the data points to their respective cluster centers. Depending on the initial estimation of the cluster centers, the resulting partitioning may be suboptimal, but the centers of the partitions will be densely populated nevertheless.

The fuzzy c-means algorithm [14] is a fuzzy extension of the k-means algorithm, which results in a fuzzy partitioning of the data set, where every data point belongs to every cluster to a certain degree (i.e., a value in [0,1]). Furthermore, the fuzzy sets generated by the fuzzy c-means algorithm are normalized in the sense that for each data point the sum of the membership degrees for the different clusters equals 1. The algorithm aims at minimizing the objective function

$$\sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\|x_i - c_j\|^2$$

where N is the number of data points, x_i is the ith data point, C is the number of clusters, c_j is the center of cluster j, and u_{ij} is the degree of membership of x_i in the cluster j. The fuzziness parameter m is an arbitrary real number greater than 1. For $m = 1$, the algorithm coincides with crisp k-means clustering. The higher the value of m, the "fuzzier" the resulting partitioning will be.

Fig. 1 shows intervals obtained by crisp 3-means clustering and membership functions obtained by fuzzy 3-means clustering for the attribute cholesterol level of the STULONG data set. Note that for $m = 2$, the partitioning is not very fuzzy, with many data points being assigned exclusively to a single cluster. In our experiments, a fuzzy partitioning with fuzziness $m = 3$ is being used.

For each attribute range, applying the k-means and the fuzzy c-means algorithm yields a crisp and a fuzzy partitioning, respectively. In general, however, the cluster centers will be different. This difference can be kept to a minimum by using the result of the k-means algorithm as the initial estimation for the fuzzy c-means algorithm. To eliminate the problem entirely, we can simply assign every data point to the nearest fuzzy cluster center in order to obtain a new crisp partitioning, which will differ only slightly from the partitioning originally obtained from the k-means algorithm.

It is our opinion that the data-driven approach to partitioning through clustering offers a good starting point for a fair comparison of quantitative versus fuzzy rule mining. In our experiments, each of the selected quantitative attributes is replaced by three either binary or fuzzy attributes, resulting in discrete and fuzzy databases with $6 \times 3 = 18$ attributes for FAM95 and HEMAT and $5 \times 3 = 15$ attributes for STULONG.

[4]The study (STULONG) was realized at the 2nd Department of Medicine, 1st Faculty of Medicine of Charles University and Charles University Hospital, U nemocnice 2, Prague 2 (head. Prof. M. Aschermann, MD, SDr, FESC), under the supervision of Prof. F. Boudík, MD, ScD, with collaboration of M. Tomecková, M.D., Ph.D., and Ass. Prof. J. Bultas, M.D., Ph.D. The data were transferred to electronic form by the European Centre of Medical Informatics, Statistics and Epidemiology of Charles University and Academy of Sciences (head. Prof. RNDr. J. Zvárová, DrSc). The data resource is on the web page http://euromise.vse.cz/challenge2004. At present, the data analysis is supported by the grant of the Ministry of Education CR Nr LN 00B 107.

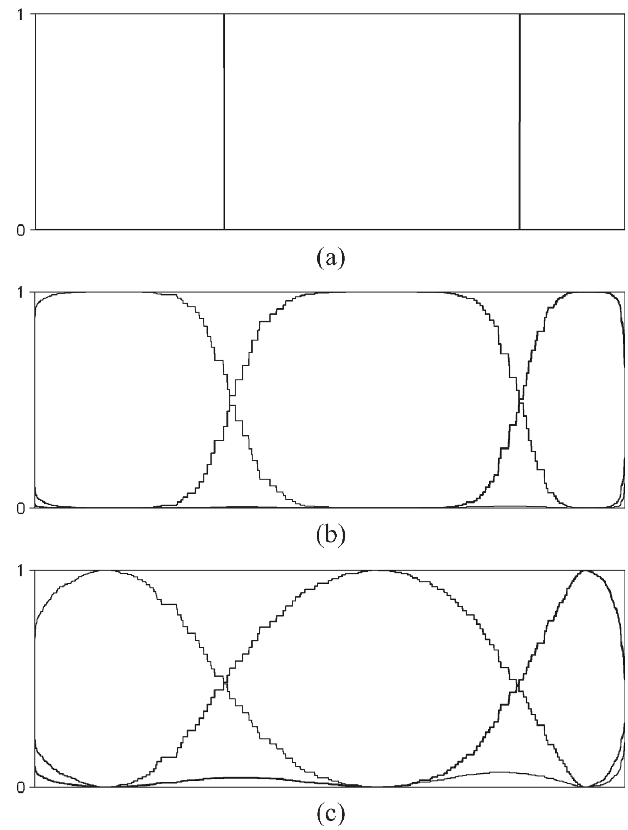

Fig. 1. Partitioning for cholesterol level. (a) Crisp $(m = 1)$. (b) Fuzzy $(m = 2)$. (c) Fuzzy $(m = 3)$.

C. Comparing Association Rules

A typical association rule mining algorithm sets a threshold for support and confidence and then selects the rules that meet both of the thresholds. The same algorithm can be used for both quantitative and fuzzy rule mining without any adaptation, because the "fuzzy definitions" of intersection, cardinality, support, and confidence coincide with the classical definitions in the binary case. In other words, whether the association rule mining algorithm is quantitative or fuzzy depends on whether the input tables constitute a crisp or a fuzzy partition of the data rather than on the algorithm itself. This should be kept in mind whenever we talk about discrete versus fuzzy algorithms.

We are interested in comparing the actual association rules obtained by quantitative and fuzzy rule mining. To keep things clear and simple, we restrict ourselves to rules with a single attribute in both the antecedent and the consequent. Such rules can be efficiently mined and are easy to understand. In a more general definition, both the antecedent and the consequent of an association rule are sets of attributes. Mining algorithms, however, tend to generate too many rules to be digested by users, and using the more general definition leads to the so-called rules explosion problem. Furthermore, there is a trend to focus on simple association rules, i.e., those containing only one attribute in the consequent, and use them as building blocks to construct more general rules if required [5], [15]. Focusing on the simplest possible association rules alleviates the need to use complicated algorithms and allows us to sort all possible rules by a certain quality measure instead of setting inherently arbitrary thresholds. The approach of ranking rules, which highly facilitates the systematic comparison of the rules, would be intractable if we allowed the inclusion of more complex rules. We get a total of $18 \times 15 = 270$ (simple) association rules for the FAM95 and the HEMAT data set (there are $6 \times 3 = 18$ attributes and each of them can be combined with each of the $5 \times 3 = 15$ attributes that refer to different underlying quantitative attributes). For the STULONG data set, we get a total of $15 \times 12 = 180$ possible rules.

We compare the rankings obtained by the quantitative and the fuzzy algorithm using the Spearman rank correlation coefficient [16] defined by

$$r_s = 1 - 6 \sum_{i=1}^{N} \frac{d_i^2}{N(N^2-1)}$$

where N is the number of ranked association rules and d_i is the difference in rank for each rule. It can be shown that $-1 \leq r_s \leq 1$ with values of r_s close to one indicating strong concordance between the two rankings and values closer to zero indicating weak concordance or independence. Negative values for r_s indicate discordance between the rankings.

D. Quantitative Versus Fuzzy Association Rules

In the first experiment, we sort 270 candidate association rules[5] by support and by confidence and compare the rankings obtained by the discrete $(m = 1)$ and the fuzzy $(m = 3)$ algorithms. We obtain a Spearman rank correlation coefficient of 0.97 for both the support and the confidence rankings, indicating a very strong agreement between the rankings.

In practical applications, only the 10 or 20 strongest rules are likely to be of interest to the user. If we compare the sets with the 20 most confident rules, the cardinality of the set difference yields 4. If we select the 20 best supported association rules, we obtain identical sets of rules, regardless of whether we use the fuzzy or the discrete algorithm!

In order to get a better feel for the actual differences, Table III lists the 20 strongest rules obtained from the discrete $(m = 1)$ and the fuzzy algorithm $(m = 3)$ along with their confidence and support values. In order to save space, we use (obvious) abbreviations for the attributes, and for clarity we also use mnemonic terms rather than the actual linguistic terms derived from cluster centers. For instance, young, middle-aged, and old actually stand for about 29, about 48, and about 73, respectively.

The rules that appear only in the left or only in the right column are in italic. We challenge the reader to compare the rankings and let him/her be the judge of which set of rules is better. In our opinion, the difference is too small to be of any real interest to the user. If we repeat these experiments for the data sets HEMAT and STULONG, we get similar results. For the support rankings, we obtain Spearman coefficients of 0.97 and 0.99, respectively. The Spearman coefficients for the confidence rankings are 0.95 and 0.98, respectively. If we only consider the 20 strongest rules, we never get more than four different rules for the fuzzy versus the quantitative approach, regardless of whether we sort rules by support or confidence.

[5] Unless explicitly stated otherwise, all experiments are based on the FAM95 data set.

E. Influence of T-Norm Operator

So far, we have assumed the use of the Zadeh intersection of fuzzy sets in the definitions of support and confidence. The underlying minimum operator is easy from a computational point of view but leaves no room for compensation. Indeed, for all x and y in [0,1], $\min(x, y) = x$ as long as $y \geq x$, regardless of the precise value of y. The work of Schweizer and Sklar [17] in probabilistic metric spaces gave a big impulse for the generalization to triangular norms that do not necessarily have this behavior.

A triangular norm (t-norm) T is a $[0,1]^2 - [0,1]$ mapping that is symmetric, associative, and nondecreasing in each argument, with $T(x, 1) = x$ for all x in [0,1]. All t-norms coincide with the classical definition of conjunction in case both arguments are binary values, but may yield very different results otherwise. The t-norms most commonly used in fuzzy applications are the minimum t-norm, the product t-norm, and the Łukasiewicz t-norm, which, respectively, lead to the following definitions of intersection of fuzzy sets A and B, i.e.,

$$\begin{aligned} (A \cap B)(x) &= \min(A(x), B(x)) \\ (A \cap B)(x) &= A(x) \cdot B(x), \qquad \text{for } x \text{ in } X \\ (A \cap B)(x) &= \max(0, A(x) + B(x) - 1). \end{aligned}$$

An interesting question is whether the difference between quantitative and fuzzy association rule mining will be bigger if we use the product t-norm or the Łukasiewicz t-norm instead of the minimum t-norm as the underlying operator for the definition of intersection. Again, however, the results show a strong agreement between the discrete and the fuzzy algorithm (Spearman coefficients of 0.98 for both the support and the confidence rankings and for both t-norms). If we only consider the strongest rules, all differences disappear, irrespective of which t-norm is being used.

Instead of comparing the discrete and the fuzzy approach, we can also compare two fuzzy approaches using different t-norms. Experiments show that in that case the Spearman coefficients are even higher (exceeding 0.99 in most cases).

IV. Conclusion

The typical argumentation or motivation for involving fuzzy set theory in association rule mining is as follows:

1) that it allows for the rules to be formulated using vague linguistic expressions, hence easier to grasp by humans;
2) that it suppresses the unwanted effect that boundary cases might cause.

The argument that rules generated by fuzzy association rule mining are more understandable to a human is not convincing in this matter because quantitative association rule mining also gives (the same strong) rules formulated in the same way in natural language. Whether these natural language expressions are represented by intervals or fuzzy sets is purely an internal matter of the system which should be of no further interest to the user.

The second argument does not consider the fact that association rule mining is not developed for toy problems but for large databases. The sharp boundary problem is already inherently suppressed and can be further minimized by using sensible partitioning methods, as is already being done in quantitative association rule mining.

Our experiments actually show that there is no significant difference between fuzzy rule mining and quantitative rule mining in large databases when using a suitable data-driven approach for attribute

TABLE III
MOST CONFIDENT RULES OBTAINED FROM THE DISCRETE AND THE FUZZY ALGORITHM. (a) DISCRETE ($m = 1$). (b) FUZZY ($m = 3$)

(a)

confidence	rule	support
.990	old age → few children	.230
.929	few persons → few children	.578
.901	low education → low hincome	.145
.887	low income → low hincome	.509
.887	old age → few persons	.206
.870	few children → few persons	.578
.850	low education → low income	.137
.848	many children → many persons	.058
.843	low hincome → low income	.509
.798	high hincome → high income	.045
.794	old age → low hincome	.185
.755	low education → few children	.122
.755	average children → average persons	.202
.750	*low income → few persons*	.431
.748	old age → low income	.174
.726	low income → few children	.417
.717	*high hincome → high education*	.041
.710	medium hincome → medium income	.241
.708	*medium income → medium hincome*	.241
.708	*average persons → average children*	.202

(b)

confidence	rule	support
.983	old age → few children	.234
.967	few persons → few children	.520
.897	low education → low hincome	.143
.884	low income → low hincome	.500
.850	low education → low income	.136
.837	low hincome → low income	.500
.808	low education → few children	.129
.801	many children → many persons	.060
.798	old age → low hincome	.190
.790	old age → few persons	.188
.784	low income → few children	.444
.768	*low hincome → few children*	.459
.766	high hincome → high income	.050
.756	old age → low income	.180
.743	average children → average persons	.151
.728	*few persons → low income*	.391
.724	*high education → few children*	.197
.720	few children → few persons	.520
.719	medium hincome → medium income	.243
.719	*medium education → few children*	.409

partitioning. In fact, in real applications, the net difference is very likely to be too small to really justify the fuzzy approach. Even when implementation issues are not taken into consideration, introducing the theoretical framework of fuzzy sets into association rule mining does not make sense unless the generated association rules are significantly different from the rules obtained by discrete mining. Our experiments show that this is certainly not generally the case.

In the experiments, we have restricted ourselves to association rules with one item in both the antecedent and the consequent. It might be interesting to see what the effect would be if we allowed more complicated rules. Also, we have only taken the two quality measures into account that were proposed by Agrawal *et al.* [12], namely, support and confidence. Association rules can be rated by a number of other quality measures as well. For a recent overview, we refer to [18]. Repeating our experiments using other quality measures may yield different results.

As a final comment, the membership functions obtained from clustering may not correspond with the most intuitive human perception of concept. Hence, we may expect rules obtained using a data-driven approach to be significantly different from the rules obtained using an expert-driven approach. The comparison of fuzzy and quantitative association rules using an expert-driven approach (for large databases) is certainly an interesting topic for future research. In this case, however, experts should also define the crisp intervals that correspond best to human intuition! The common practice of comparing data-driven crisp data mining with expert-driven fuzzy data mining does not provide convincing arguments for the introduction of fuzzy association rules.

ACKNOWLEDGMENT

The authors would like to thank the anonymous referees for their constructive suggestions.

REFERENCES

[1] P. Hájek and T. Havránek, *Mechanizing Hypothesis Formation: Mathematical Foundations for a General Theory*. Berlin, Germany: Springer-Verlag, 1987.

[2] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proc. ACM SIGMOD Int. Conf. Management of Data*, Washington, DC, 1993, pp. 207–216.

[3] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," in *Proc. ACM SIGMOD Conf. Management Data*, Montréal, QC, Canada, 1996, pp. 1–12.

[4] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–358, Jun. 1965.

[5] G. Chen, P. Yan, and E. E. Kerre, "Computationally efficient mining for fuzzy implication-based association rules in quantitative databases," *Int. J. Gen. Syst.*, vol. 33, no. 2–3, pp. 163–182, Apr./Jun. 2004.

[6] M. De Cock, C. Cornelis, and E. E. Kerre, "Fuzzy association rules: A two-sided approach," in *Proc. Int. Conf. Fuzzy Information Processing—Theories and Applications*, Beijing, China, 2003, pp. 385–390.

[7] M. Delgado, N. Marín, D. Sánchez, and M.-A. Vila, "Fuzzy association rules: General model and applications," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 2, pp. 214–225, Apr. 2003.

[8] A. Gyenesei, "A fuzzy approach for mining quantitative association rules," *Acta Cybern.*, vol. 15, no. 2, pp. 305–320, Dec. 2001.

[9] T.-P. Hong, C.-S. Kuo, and S.-C. Chi, "Trade-off between computation time and number of rules for fuzzy mining from quantitative data," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 9, no. 5, pp. 587–604, Oct. 2001.

[10] E. Hüllermeier, "Mining implication-based fuzzy association rules in databases," in *Proc. 9th Int. Conf. Information Processing and Management Uncertainty Knowledge-Based Systems*, Annecy, France, 2002, pp. 101–108.

[11] C. M. Kuok, A. W.-C. Fu, and M. H. Wong, "Mining fuzzy association rules in databases," *ACM SIGMOD Rec.*, vol. 27, no. 1, pp. 41–46, Mar. 1998.

[12] A. Gyenesei and J. Teuhola, "Multidimensional fuzzy partitioning of attribute ranges for mining quantitative data," *Int. J. Intell. Syst.*, vol. 19, no. 11, pp. 1111–1126, Nov. 2004.

[13] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Mathematical Statistics and Probability*, Berkeley, CA, 1967, pp. 281–297.

[14] F. Hoppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis, Methods for Classification, Data Analysis and Image Recognition*. New York: Wiley, 1999.

[15] G. Chen, Q. Wei, D. Liu, and G. Wets, "Simple association rules (sar) and the sar-based rule discovery," *Comput. Ind. Eng.*, vol. 43, no. 4, pp. 721–733, Sep. 2002.

[16] E. L. Lehmann and H. D'Abrera, *Nonparametrics: Statistical Methods Based on Ranks*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[17] B. Schweizer and A. Sklar, "Associative functions and statistical triangle inequalities," *Publ. Math.*, vol. 8, pp. 169–186, 1961.

[18] P. N. Tan, V. Kumar, and J. Srivastava, "Selecting the right interestingness measure for association patterns," in *Proc. 8th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Edmonton, AB, Canada, 2002, pp. 32–41.