# A Novel Method for Discovering Fuzzy Sequential Patterns Using the Simple Fuzzy Partition Method

**Ruey-Shun Chen, Yi-Chung Hu**
*Institute of Information Management, National Chiao Tung University, Hsinchu 300, Taiwan, ROC.*
*E-mail address: rschen@bis03.iim.nctu.edu.tw (R. -S. Chen)*

**Sequential patterns refer to the frequently occurring patterns related to time or other sequences, and have been widely applied to solving decision problems. For example, they can help managers determine which items were bought after some items had been bought. However, since fuzzy sequential patterns described by natural language are one type of fuzzy knowledge representation, they are helpful in building a prototype fuzzy knowledge base in a business. Moreover, each fuzzy sequential pattern consisting of several fuzzy sets described by the natural language is well suited for the thinking of human subjects and will help to increase the flexibility for users in making decisions. Additionally, since the comprehensibility of fuzzy representation by human users is a criterion in designing a fuzzy system, the simple fuzzy partition method is preferable. In this method, each attribute is partitioned by its various fuzzy sets with pre-specified membership functions. The advantage of the simple fuzzy partition method is that the linguistic interpretation of each fuzzy set is easily obtained. The main aim of this paper is exactly to propose a fuzzy data mining technique to discover fuzzy sequential patterns by using the simple partition method. Two numerical examples are utilized to demonstrate the usefulness of the proposed method.**

## 1. Introduction

Data mining is the exploration and analysis of data in order to discover meaningful patterns (Berry & Linoff, 1997). Thus knowledge acquisition can be easily achieved for users by checking these patterns discovered from databases, and association rule is an important type of knowledge representation. Agrawal et al. (Agrawal, Imielinski, & Swami, 1993) initially proposed a method to find association rules, later proposing the well-known Apriori algorithm (Agrawal, Mannila, Srikant, Toivonen, &Verkamo, 1996). In addition to association rules, sequential patterns are another important type of knowledge representation, and effective algorithms (i.e. AprioriSome and AprioriAll) for mining sequential patterns were proposed by Agrawal and Srikant (1995). In addition, sequential patterns have been widely applied to solve decision problems. For example, they can help managers determine which items were bought after some items had (already) been bought (Han & Kamber, 2001), or realize browsing orders of homepages in a web site (Myra, 2000).

Sequential pattern mining is the mining of frequently occurring patterns related to time or other sequences (Han & Kamber, 2001), where a sequence is an ordered list of itemsets (Agrawal & Srikant, 1995). Specially, if there are $k$ itemsets ($k \geq 1$) in a frequent sequence whose support is larger than or equal to the user-specified minimum support, then we call it a frequent $k$-sequence. Moreover, a sequential pattern is a frequent sequence but it is not contained in another sequence (Agrawal & Srikant, 1995). For example, a 2-sequence ⟨{Banana} , {Apple, Orange}⟩ may represent items Apple and Orange being bought together after item Banana had been bought, where {Banana} and {Apple, Orange} are itemsets. Whereas ⟨{Banana} , {Apple, Orange}⟩ is not contained in the 1-sequence ⟨{Banana, Apple, Orange}⟩ since the latter sequence is shorter than the former sequence.

However, since fuzzy sequential patterns described by natural language are one type of fuzzy knowledge representation, they are helpful to build a prototype fuzzy knowledge base in business. Moreover, fuzzy sequential patterns described by the natural language are well suited for the thinking of human subjects and will help to increase the flexibility for users in making decisions. Actually, each

Nomenclature $K$, number of partitions in each quantitative attribute; $k$, length of a fuzzy sequence; $d$, degree of a given relation, where $d \geq 1$; $A^{x_m}_{K,i_m}$, $i_m$-th linguistic value of $K$ fuzzy partitions defined in quantitative attribute $x_m$, $1 \leq im \leq K$; $\mu^{x_m}_{K,i_m}$, membership function of $A^{x_m}_{K,i_m}$; $n$, total number of customers; $c_r$, $r$-th customer, where $1 \leq r \leq n$; $\alpha_r$, number of consecutive transactions ordered by transaction-time for $c_r$; $\beta$, total number of frequent fuzzy grids; $t^{(r)}_p$, $p$-th transaction corresponding to $c_r$, where $t^{(r)}_p = (t^{(r)}_{p1}, t^{(r)}_{p2}, \ldots, t^{(r)}_{pd})$, and $1 \leq p \leq \alpha_r$; $L_j$, $j$-th frequent fuzzy grid, where $1 \leq j \leq \beta$.

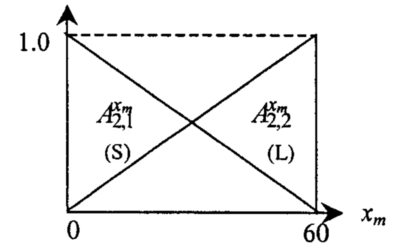


FIG. 1. $K = 2$ for $x_m$.

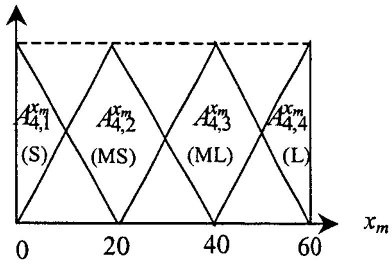


FIG. 3. $K = 4$ for $x_m$.

fuzzy sequential pattern is composed of several fuzzy sets that can be described by the natural language. For example, a fuzzy sequence $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle$ discovered from a transaction database in summer may represent that large purchase amounts of product 2 were bought by customers after they had bought small purchase amounts of product 1, where $A^{\text{Product 1}}_{2,1}$ and $A^{\text{Product 2}}_{2,2}$ are fuzzy sets defined in purchase amounts of product 1 (i.e., Product 1) and purchase amounts of product 2 (i.e., Product 2), respectively. In other words, a fuzzy sequence expresses the temporal relation between purchase behaviors described by fuzzy sets. In addition, if one customer bought product 1 on August 1, and bought product 2 on August 6, then we say that the corresponding transaction record supports $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle$.

Since the comprehensibility of fuzzy representation by human users is a criterion in designing a fuzzy knowledge-based system (Ishibuchi et al., 1999), easy linguistic interpretations of fuzzy sets must be taken into account. The simple fuzzy partition method is thus preferable (Ishibuchi et al., 1999). In this method, each attribute is partitioned by its various fuzzy sets with pre-specified membership functions. For example, the $m$-th axis denoted by $x_m$ is partitioned by 2, 3, 4 linguistic values with linguistic interpretations as shown in Figures 1–3, respectively. For example, $A^{x_1}_{3,1}$ and $A^{x_2}_{3,3}$ are interpreted as "small" and "large", respectively. That is, the simple fuzzy partition method provides a comprehensible expression for interpreting fuzzy sets.

The main aim of this paper is to propose a fuzzy data mining technique to discover fuzzy sequential patterns by using the simple partition method. The first phase is to find purchase behaviors (e.g., $A^{\text{Product 1}}_{2,1}$, $A^{\text{Product 2}}_{2,2}$) that frequently occurred for a period of time, and the second phase is to discover fuzzy sequential patterns by analyzing the temporal relation between those purchase behaviors (e.g., $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle$ or $\langle A^{\text{Product 2}}_{2,2}, A^{\text{Product 1}}_{2,1}\rangle$) found in the first phase.

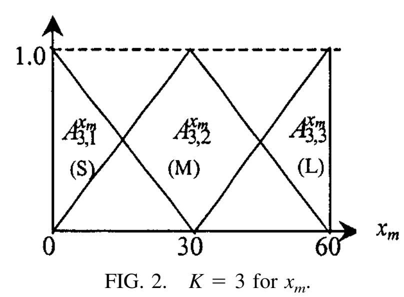


FIG. 2. $K = 3$ for $x_m$.

The rest of this paper is organized as follows. The simple fuzzy partition method is introduced in Section 2. The determinations of purchase behaviors that frequently occurred (i.e., frequent fuzzy grids) and fuzzy sequential patterns are presented in Section 3 in detail. The framework of the proposed method consisting of two phases is also illustrated in this section. The real implementation of the proposed method is described in Section 4. In Section 5, two numerical examples are used to demonstrate the usefulness of the proposed method. We end this paper with discussions and conclusions in Section 6.

## 2. Simple Fuzzy Partition Method

The concepts of a linguistic variable were proposed by Zadeh (1976), who initially proposed the fuzzy sets (Zadeh, 1965). A linguistic variable is a variable whose values are linguistic terms or sentences in a natural language. We view both quantitative and categorical attributes, which are used to describe each sample data, as linguistic variables (Zadeh, 1976; Chen & Jong, 1997; Zimmermann, 1996; Pedrycz & Gomide, 1994). For example, the values of the linguistic variable "Purchase amount of Product 1" may be "medium" or "very close to 50." Here, "medium" or "very close to 50" are called linguistic values. Then, each linguistic variable can be partitioned by its linguistic values with pre-specified membership functions, so-called the simple fuzzy partition method. Simple fuzzy grids or grid partitions in a pattern space (Ishibuchi , Nozaki, Yamamoto, & Tanaka, 1995; Ishibuchi et al, 1999; Jang & Sun, 1995) are thus generated.

The simple fuzzy partition method has been widely used in pattern recognition and fuzzy reasoning. For example, there are the applications to classification rule discovery for pattern classification problems (Ishibuchi et al, 1995; Ishibuchi et al, 2001; Hu, Tzeng, & Chen, 2002; Ishibuchi et al, 1999; Ravi & Zimmermann, 2000), and to the fuzzy rule generation for control problems (Wang & Mendel, 1992; Homaifar & McCormick, 1995; Jang, 1993). In addition, several fuzzy approaches for partitioning a pattern space were discussed by Sun (1994) and Bezdek (1981). From the above-mentioned studies, we can find that it should be quite feasible to discover useful fuzzy knowledge from business databases by utilizing the simple fuzzy partition method. Moreover, as we have mentioned above, the simple fuzzy partition method provides a comprehensible expression to interpret fuzzy sets. Additionally, data mining techniques

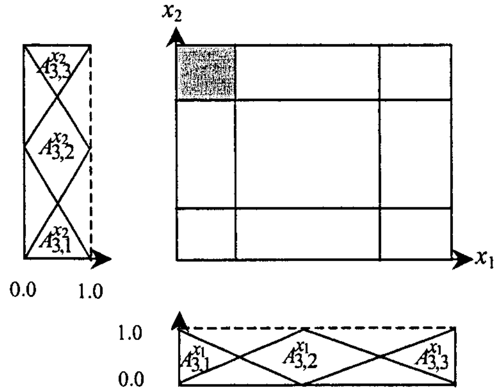


FIG. 4. Both attributes $x_1$ and $x_2$ are partitioned into three partitions.

can ease the knowledge acquisition bottleneck in building prototype knowledge base systems (Hong, Wang, Wang, & Chien, 2000). Fuzzy data mining techniques using the simple fuzzy partition method are thus helpful to build a comprehensibly prototype fuzzy knowledge base in a business.

In the simple fuzzy partition method, $K$ ($K = 2, 3, 4, \ldots$) various linguistic values are defined in each quantitative attribute. $K$ is also pre-specified before executing the proposed method. It is clear that if $K$ is very small (e.g., $K = 2$), then the resultant partition is too coarse; otherwise (e.g., $K = 10$), the resultant partition is too fine. It is also suggested that $K$ should not exceed 9, since it is difficult for us to simultaneously judge or distinguish from the items whose number is larger than 9 (Ravi & Zimmermann, 2000).

Triangular membership functions, which are usually used for the linguistic values, are the default value in our method. In fact, Pedrycz (1994) had pointed out the usefulness and effectiveness of the triangular membership functions in the fuzzy modeling. It is noted that, in addition to the simple fuzzy partition method, decision-makers can also specify membership functions and the number of partitions in each attribute depending on professional knowledge they have or subject perception.

Specifically, we use $A^{x_m}_{K,i_m}$ to denote a candidate 1-dim fuzzy grid. And, $\mu^{x_m}_{K,i_m}(x)$ is defined as follows:

$$\mu^{x_m}_{K,i_m}(x) = \max\{1 - |x - a^K_{i_m}|/b^K, 0\} \quad (1)$$

where

$$a^K_{i_m} = mi + (ma - mi) \cdot (i_m - 1)/(K - 1) \quad (2)$$

$$b^K = (ma - mi)/(K - 1) \quad (3)$$

where $ma$ is the maximum value of attribute domain, and $mi$ is the minimum value.

If we partition both $x_1$ and $x_2$ into three fuzzy partitions, then a pattern space is divided into $3 \times 3$ 2-dim fuzzy grids, as shown in Figure 4. For the shaded 2-dim fuzzy grid shown in Figure 4, we use linguistic values, $A^{x_1}_{3,1} \times A^{x_2}_{3,3}$ (i.e. small AND large), to represent it. In this paper, each fuzzy grid is treated as a purchase behavior. The next important task is how to use these candidate 1-dim fuzzy grids to generate frequent fuzzy sequences and fuzzy sequential patterns. The framework of the proposed method is further described in following section.

## 3. Determine Frequent Fuzzy Grids and Fuzzy Sequential Patterns

In this section, the concrete meanings of frequent fuzzy grids and fuzzy sequential patterns are described in detail. At first, the computational steps of the proposed method are briefly introduced as follows.

After candidate 1-dim fuzzy grids have been generated, we must determine how to find frequent fuzzy grids, frequent fuzzy $k$-sequences ($k \geq 1$) and fuzzy sequential patterns from those candidate 1-dim fuzzy grids. Frequent fuzzy grids with a small dimension, say $m$, are used to construct candidate ($m + 1$)-dim fuzzy grids accompanied by a fuzzy support. A candidate ($m + 1$)-dim grid can be determined to be frequent or not by comparing its fuzzy support with the user-specified minimum fuzzy support (min FS). At the end of phase I, each frequent fuzzy grid, say $L_j$, can be transformed into a frequent fuzzy 1-sequence $\mathrm{L_j}$. Frequent fuzzy grid may stand for purchase behaviors that frequently occurred for a period of time.

We define that a fuzzy sequence is an ordered list of frequent fuzzy grids, and the length of a fuzzy sequence is the number of frequent fuzzy grids in the fuzzy sequence. That is, a fuzzy sequence expresses the temporal relation between frequent fuzzy grids. Thus, if there are $k$ fuzzy grids ($k \geq 1$) in a fuzzy sequence, then we call it a fuzzy $k$-sequence. For example, $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2} \rangle$ is a fuzzy 2-sequence.

The main purpose of phase II is to discover fuzzy sequential patterns by analyzing the temporal relation between those frequent grids found in phase I. In phase II, frequent fuzzy sequences with a shorter length, say $k$, are used to construct candidate fuzzy sequences with a longer length (i.e. fuzzy ($k + 1$)-sequences) accompanied by a fuzzy support. A candidate fuzzy ($k$+1)-sequence can also be determined to be frequent or not by comparing its fuzzy support with the min FS used in phase I. At the end of phase II, all fuzzy sequential patterns are generated from those frequent fuzzy sequences.

From the above-mentioned operations, the framework for discovering fuzzy sequential patterns is illustrated in Figure 5. Below, the Apriori algorithm is briefly introduced in Subsection 3.1 since the concept of support is originated from this well-known algorithm. The determinations of frequent fuzzy 1-sequences and fuzzy sequential patterns are described in Subsections 3.2 and 3.3, respectively.

### *3.1 The Apriori Algorithm*

Association rules are one type of knowledge representation, having been widely applied to analyze market baskets

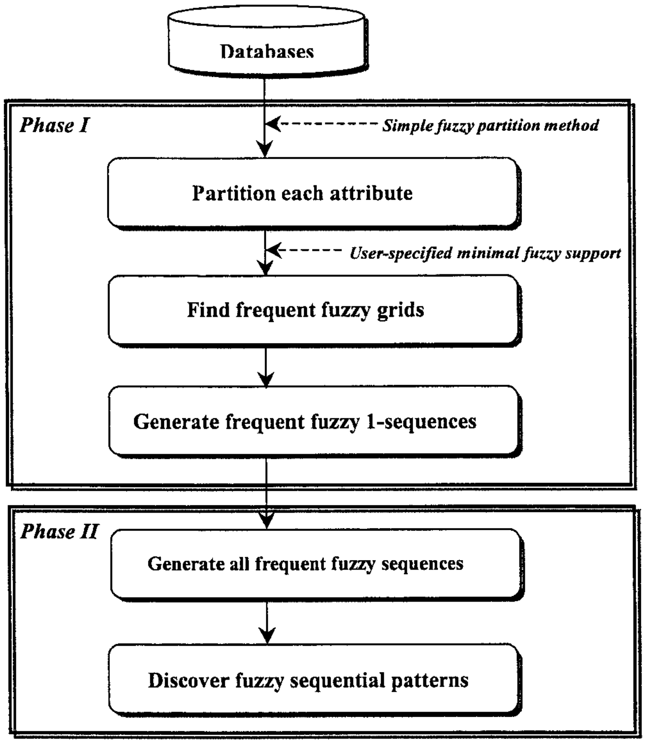


FIG. 5. Framework of the proposed method.

to help managers realize which items are likely to be bought at the same time (Han & Kamber, 2001, Berry & Linoff, 1997). The Apriori algorithm proposed by Agrawal et al. (1996) is an influential algorithm that can be used to find association rules. In this algorithm, all frequent itemsets are found from databases in the first phase. A candidate itemset is frequent if its support is larger than or equal to the user-specified minimum support. Generally, a frequent itemset means that this set is useful for decision makers. In the second phase, frequent itemsets are used to generate association rules.

**Example 1.** Any set of fruits, say {Apple, Orange}, sold in one supermarket is a candidate itemset. The support of {Apple, Orange} is computed by dividing the number of transactions that buy apples and oranges by the total number of transactions. That is, if Apple and Orange were bought at the same time in one transaction, then this transaction record supports {Apple, Orange}. If the support of {Apple, Orange} is larger than or equal to the user-specified minimum support, then {Apple, Orange} is a frequent itemset. This means that "Apple and Orange are likely to be bought together" frequently occurred.

Furthermore, the larger minimum support is specified by users, the smaller number of frequent itemsets will be generated. Therefore, if the user-specified minimum support is set to zero, then all itemsets will be frequent. Additionally, the Apriori property (Han & Kamber, 2001) for mining association rules shows that any subset of a frequent itemset must also be frequent. This property can be also applied to show that any superset (i.e., sequence of length above $k$) of an infrequent $k$-sequence is not frequent (Han & Kamber, 2001).

**Example 2.** From the Apriori property, if {Banana, Apple, Orange} is frequent, then {Banana} , {Apple} , {Orange} , {Banana, Apple} , {Banana, Orange} and {Apple, Orange} must be frequent. ⟨{Banana} , {Apple, Orange}⟩ is an infrequent 2-sequence if either ⟨{Banana}⟩ or {Apple, Orange} is infrequent.

### *3.2 Frequent Fuzzy 1-Sequences*

In phase I, the main work is to generate frequent fuzzy grids, and then transform those grids to frequent fuzzy 1-sequences. Now, given a candidate $l$-dim ($l \leq d$) fuzzy grid $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$, the degree to which $t^{(r)}_p$ belongs to this fuzzy grid can be computed as follows (Ishibuchi et al., 2001; Hu, Chen, & Tz, 2002) :

$$\mu_{A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}}(t^{(r)}_p) = \mu^{x_1}_{K,i_1}(t^{(r)}_{P_1}) \cdot \mu^{x_2}_{K,i_2}(t^{(r)}_{P_2}) \cdot \cdots \cdot \mu^{x_{l-1}}_{K,i_{l-1}}(t^{(r)}_{P_{l-1}}) \cdot \mu^{x_l}_{K,i_l}(t^{(r)}_{P_l}) \quad (7)$$

where " · " is a fuzzy intersection operator, namely the algebraic product (Zimmermann, 1996; Hu et al., 2002). It should be noted that in comparison with the other fuzzy intersection operators such as the minimum operator or the drastic product, the algebraic product "gently" performs the fuzzy intersection. To check whether this fuzzy grid is frequent or not, we define its fuzzy support, $\mathrm{FS}(A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l})$, as follows:

$$\begin{aligned} &\mathrm{FS}(A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}) \\ &= \sum_{r=1}^{n} \mu_{A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}}(c_r)/n \\ &= \sum_{r=1}^{n} \max_{p=1 \cdots a_r} \left[ \sum_{r=1}^{n} \mu_{A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}}(t^{(r)}_p) \right] \Big/ n \end{aligned} \quad (8)$$

where $\mu_{A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}}(c_r)$ is the degree to which $c_r$ supports $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$. Since sequential pattern mining mainly analyzes the customer behaviors, the fuzzy support is obtained by computing $\mu_{A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_2}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}}(c_r)$. If $\mathrm{FS}(A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l})$ is larger than or equal to the user-specified minimum fuzzy support (i.e. min FS), then $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$ is a frequent $l$-dim fuzzy grid. The fuzzy support also indicates the degree of importance of one fuzzy grid. Of course, if the user-specified minimum fuzzy support is set to zero, then all $l$-dim fuzzy grid ($1 \leq l \leq d$) will be frequent. Actually, $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$ is a fuzzy subset and can be represented as a Zadeh fuzzy notation (Zimmermann, 1996; Pedrycz & Gomide, 1998):

$$A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l} = \sum_{r=1}^{n} \mu_{A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}}(c_r)/c_r \quad (9)$$

It is clear that $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_{l-1}} \subseteq A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}}$ holds. From Eqs. (8) and (9), we can observe that if a $m$-dim fuzzy grid, say $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{m-1}}_{K,i_{m-1}} \times A^{x_m}_{K,i_m}$, that participates in the construction of a frequent $l$-dim fuzzy grid, say $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{m-1}}_{K,i_{m-1}} \times A^{x_m}_{K,i_m} \times A^{x_{m+1}}_{K,i_{m+1}} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$, then that $m$-dim fuzzy grid must also be frequent since min FS $\leq$ FS$(A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{m-1}}_{K,i_{m-1}} \times A^{x_m}_{K,i_m} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}) \leq$ FS$(A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{m-1}}_{K,i_{m-1}} \times A^{x_m}_{K,i_m})$ holds. That is, any subset of a frequent fuzzy grid must also be frequent. Finally, $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$ can be transformed into a frequent fuzzy 1-sequence denoted by $\langle A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$.

**Example 3.** Any two fuzzy grids in $\{A^{x_1}_{3,2} \times A^{x_2}_{3,1}, A^{x_1}_{3,2} \times A^{x_3}_{3,3}, A^{x_2}_{3,1} \times A^{x_3}_{3,3}\}$ can be used to generate $A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ since $A^{x_1}_{3,2} \times A^{x_2}_{3,1} \subseteq A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$, $A^{x_1}_{3,2} \times A^{x_3}_{3,3} \subseteq A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$, and $A^{x_2}_{3,1} \times A^{x_3}_{3,3} \subseteq A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ hold. If $A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ is frequent, then $A^{x_1}_{3,2} \times A^{x_2}_{3,1}$, $A^{x_1}_{3,2} \times A^{x_3}_{3,3}$ and $A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ are also frequent. Therefore, the above-mentioned fuzzy grids will be transformed into $\langle A^{x_1}_{3,2} \times A^{x_2}_{3,1}\rangle$, $\langle A^{x_1}_{3,2} \times A^{x_3}_{3,3}\rangle$, $\langle A^{x_2}_{3,1} \times A^{x_3}_{3,3}\rangle$, and $\langle A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}\rangle$, respectively.

### *3.3 Fuzzy Sequential Patterns*

Based on frequent fuzzy 1-sequence found in phase I, the next step for us is to find frequent fuzzy $k$-sequences ($2 \leq k \leq \beta$). As we have mentioned in the previous section, each frequent fuzzy grid, say $L_j$, can be transformed into a frequent fuzzy 1-sequence denoted by $L_j$. The fuzzy support of a fuzzy $k$-sequence is the average degree of total customers who support this sequence. Here, we take a fuzzy $k$-sequence $L_1, L_2, \ldots, L_k$, which may represent $L_1, L_2, \ldots, L_k$ being bought sequentially, to be an example to compute its fuzzy support. For the $r$-th customer (i.e., $c_r$) with $\alpha_r$ transactions, there are $\alpha_r\, C_k$ ($\alpha_r \geq k$) different combinations $(t^{(r)}_{s_1}, t^{(r)}_{s_2}, \ldots, t^{(r)}_{s_k})$ ($1 \geq s_1 < s_2 < \cdots < s_k \leq \alpha_r$) ordered by transaction-time. Since $(t^{(r)}_{s_1}, t^{(r)}_{s_2}, \cdots t^{(r)}_{s_k})$ supports $\langle L_1, L_2, \ldots, L_k\rangle$, the degree FS$(\langle L_1, L_2, \ldots, L_k\rangle_r)$ to which $c_r$ supports $L_1, L_2, \ldots, L_k$ is described as follows:

$$\text{FS}(\langle L_1, L_2, \ldots, L_k\rangle_r) = \max_{(t^{(r)}_{s1}, t^{(r)}_{s2}, \ldots, t^{(r)}_{sk})} [\mu_{L1}(t^{(r)}_{s1}) \cdot \mu_{L2}(t^{(r)}_{s2}) \cdot \cdots \cdot \mu_{Lk}(t^{(r)}_{sk})],$$
$$\text{for } \alpha_r C_k \text{ different } (t^{(r)}_{s1}, t^{(r)}_{s2}, \ldots, t^{(r)}_{sk}) \quad (10)$$

where $\mu_{Lk}(t^{(r)}_{s_k})$ represents the degree which $t^{(r)}_{s_k}$ belongs to $L_k$, and can be computed by Eq. (7). Of course, if $\alpha_r < k$, then FS$\langle L_1, L_2, \ldots, L_k\rangle_r = 0$.

**Example 4.** Assume that the number of transactions of $c_2$ is $\alpha_2 = 3$ (i.e., $t^{(2)}_1$, $t^{(2)}_2$, and $t^{(2)}_3$), and all possible combinations of transactions ordered by transaction-time is $(t^{(2)}_1, t^{(2)}_2)$, $(t^{(2)}_1, t^{(2)}_3)$, and $(t^{(2)}_2, t^{(2)}_3)$. Let $L_1$ and $L_2$ be $A^{\text{Product 2}}_{3,2}$ and $A^{\text{Product 3}}_{3,2}$, respectively, FS$(\langle A^{\text{Product 2}}_{3,2}, A^{\text{Product 3}}_{3,2}\rangle)_2$ (i.e., $k = 2$) is obtained by computing $\max\{\mu_{L1}(t^{(2)}_1) \cdot \mu_{L2}(t^{(2)}_2),\ \mu_{L1}(t^{(2)}_1) \cdot \mu_{L2}(t^{(2)}_3),\ \mu_{L1}(t^{(2)}_2) \cdot (t^{(2)}_3)\}$ since each combination may support $\langle A^{\text{Product 2}}_{3,2}, A^{\text{Product 3}}_{3,2}\rangle$ and $\alpha_2 \geq$ k. However, if the number of transactions of $c_1$ is $\alpha_1 = 1$, then FS$(\langle A^{\text{Product 2}}_{3,2}, A^{\text{Product 3}}_{3,2}\rangle)_1$ is equal to zero since $\alpha_1 < k$.

The fuzzy support FS$(\langle L_1, L_2, \ldots, L_k\rangle)$ of $\langle L_1, L_2, \ldots, L_k\rangle$ is further described as follows:

$$\text{FS}(\langle L_1, L_2, \cdots, L_k\rangle) = \sum_{r=1}^{n} \text{FS}(\langle L_1, L_2, \cdots, L_k\rangle_r)/n \quad (11)$$

**Example 5.** Following Example 4, if the total number of customers is 2 (i.e., $n = 2$), then FS$(\langle A^{\text{Product 2}}_{3,2}, A^{\text{Product 3}}_{3,2}\rangle) =$ [FS$(\langle A^{\text{Product 2}}_{3,2}, A^{\text{Product 3}}_{3,2}\rangle)_1$ + FS$(\langle A^{\text{Product 2}}_{3,2}, A^{\text{Product 3}}_{3,2}\rangle)_2$]/2. The fuzzy support also indicates the degree of importance of one fuzzy sequence.

If FS$\langle L_1, L_2, \ldots, L_k\rangle$ is larger than or equal to the aforementioned min FS, then $\langle L_1, L_2, \ldots, L_k\rangle$ is a frequent fuzzy $k$-sequence. From Eqs. (10) and (11), it is clear that FS$(\langle L_1, L_2, \ldots, L_{k+1}\rangle) \leq$ FS$(\langle L_1, L_2, \ldots, L_k\rangle)$ since FS$(\langle L_1, L_2, \ldots, L_{k+1}\rangle_r) \leq$ FS$(\langle L_1, L_2, \ldots, L_k\rangle_r)$ ($1 \leq r \leq n$) holds. In addition, any fuzzy ($k$+1)-sequence, say $\langle L_1, L_2, \ldots, L_{k+1}\rangle$, cannot be frequent if fuzzy $k$-sequence $\langle L_1, L_2, \ldots, L_k\rangle$ is infrequent since FS$(\langle L_1, L_2, \ldots, L_{k+1}\rangle) \leq$ FS$(\langle L_1, L_2, \ldots, L_k\rangle) \leq$ min FS holds.

**Example 6.** $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle$ is not frequent if either $\langle A^{\text{Product 1}}_{2,1}\rangle$ or $\langle A^{\text{Product 2}}_{2,2}\rangle$ is not frequent since FS$(\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle_r) \leq$ FS$(A^{\text{Product 1}}_{2,1}\rangle_r) \leq$ min FS and FS$(\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle_r) \leq$ FS$(\langle A^{\text{Product 2}}_{2,2}\rangle_r) \leq$ min FS hold. In other words, if $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 2}}_{2,2}\rangle$ is a frequent sequence, then $\langle A^{\text{Product 1}}_{2,1}\rangle$ and $\langle A^{\text{Product 2}}_{2,2}\rangle$ are frequent.

As for $\langle A^{x_1}_{2,1}, A^{x_1}_{2,2}, A^{x_1}_{2,1} \times A^{x_1}_{2,2}\rangle$, it can be generated by using any two fuzzy sequences in $\{\langle A^{x_1}_{2,1}, A^{x_1}_{2,2}\rangle, \langle A^{x_1}_{2,1}, A^{x_1}_{2,1} \times A^{x_1}_{2,2}\rangle, \langle A^{x_1}_{2,2}, A^{x_1}_{2,1} \times A^{x_1}_{2,2}\rangle\}$. If $\langle A^{x_1}_{2,1}, A^{x_1}_{2,2}, A^{x_1}_{2,1} \times A^{x_1}_{2,2}\rangle$ is frequent, then $\langle A^{x_1}_{2,1}, A^{x_1}_{2,2}\rangle$, $\langle A^{x_1}_{2,1}, A^{x_1}_{2,1} \times A^{x_1}_{2,2}\rangle$, and $\langle A^{x_1}_{2,2}, A^{x_1}_{2,1} \times A^{x_1}_{2,2}\rangle$ must be also frequent.

As we have mentioned above, a sequential pattern is a frequent sequence but it is not contained in another sequence. Formally, a frequent $z1$-sequence, say $a$, denoted by $\langle L_{a,1}, L_{a,2}, \ldots, L_{a,z1}\rangle$ is contained in another frequent $z2$-sequence, say $b$, denoted by $\langle L_{b,1}, L_{b,2}, \ldots, L_{b,z2}\rangle$, if $z1 \leq z2$ and there exist integers $1 \leq j1 < j2 < \ldots < jz1 \leq z2$ such that $L_{a,1} \subseteq L_{b,j1}, L_{a,2} \subseteq L_{b,j2}, \ldots, L_{a,z1} \subseteq L_{b,jz1}$. Then, $a$ is not a sequential pattern but $b$ is if it is not contained in the other frequent sequences. A detailed example is shown as follows.

**Example 7.** The 2-sequence ⟨{Banana} , {Apple}⟩ denoted by $a$ is contained in ⟨{Banana} , {Apple, Orange}⟩ since {Banana} ⊆ {Banana} and {Apple} ⊆ {Apple, Orange} . Because {Banana} ⊆ {Banana} and {Apple} ⊆ {Apple,} , ⟨{Banana} , {Apple}⟩ is also contained in ⟨{Ba-

TABLE 1. An initial table FGTTFS for an example.

| Fuzzy grid | FG | | | | TT | | FS |
|---|---|---|---|---|---|---|---|
| | $A^{x1}_{2,1}$ | $A^{x1}_{2,2}$ | $A^{x2}_{2,1}$ | $A^{x2}_{2,2}$ | $t^{(1)}_1$ | $t^{(1)}_2$ | |
| $A^{x1}_{2,1}$ | 1 | 0 | 0 | 0 | $\mu^{x1}_{2,1}(t^{(1)}_{1_1})$ | $\mu^{x1}_{2,1}(t^{(1)}_{2_1})$ | $\text{FS}(A^{x1}_{2,1})$ |
| $A^{x1}_{2,2}$ | 0 | 1 | 0 | 0 | $\mu^{x1}_{2,2}(t^{(1)}_{1_1})$ | $\mu^{x1}_{2,2}(t^{(1)}_{2_1})$ | $\text{FS}(A^{x1}_{2,2})$ |
| $A^{x2}_{2,1}$ | 0 | 0 | 1 | 0 | $\mu^{x2}_{2,1}(t^{(1)}_{1_2})$ | $\mu^{x2}_{2,1}(t^{(1)}_{2_2})$ | $\text{FS}(A^{x2}_{2,1})$ |
| $A^{x2}_{2,2}$ | 0 | 0 | 0 | 1 | $\mu^{x2}_{2,2}(t^{(1)}_{1_2})$ | $\mu^{x2}_{2,2}(t^{(1)}_{2_2})$ | $\text{FS}(A^{x2}_{2,2})$ |

nana} , {Apple} , {Orange}⟩. From the aforementioned definition, we can find that $a$ is not contained in the 1-sequence ⟨{Banana, Apple, Orange}⟩ denoted by $b$ since the latter sequence (i.e., $b$ with length 1) is shorter than the former sequence (i.e., $a$ with length 2).

Similarly, we define that a fuzzy sequential pattern is a frequent fuzzy sequence, but it is not contained in any other frequent fuzzy sequence. That is, not all frequent fuzzy sequences are desirable. For measuring the importance of a fuzzy sequence, in addition to the fuzzy support, the amount or type of information contained in it is the other criterion. Formally, a frequent fuzzy $z1$-sequence, say $a$, denoted by $\langle L_{a,1}, L_{a,2}, \ldots, L_{a,z1}\rangle$ is contained in another frequent fuzzy $z2$-sequence, say $b$, denoted by $\langle L_{b,1}, L_{b,2}, \ldots, L_{b,z2}\rangle$, if $z1 \leq z2$ and there exist integers $1 \leq j1 \leq j2 < \cdots < jz1 \leq z2$ such that $L_{b,j1} \subseteq L_{a,1}, L_{b,j2} \subseteq L_{a,2}, \ldots, L_{b,jz1} \subseteq L_{a,z1}$. Then, $a$ is not a fuzzy sequential pattern but $b$ is if it is not contained in the other frequent fuzzy sequences. In comparison with $b$, it seems that $a$ is not valuable for decision makers.

**Example 8.** Assume that $\langle L_{a,1}, L_{a,2}\rangle = \langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2}\rangle$ (i.e., $z1 = 2$) and $\langle L_{b,1}, L_{b,2}, L_{b,3}\rangle = \langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2} \times A^{\text{Product 2}}_{2,1}, A^{\text{Product 1}}_{2,1} \times A^{\text{Product 2}}_{2,2}\rangle$ (i.e., $z2 = 3$) are frequent fuzzy sequences. $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2}\rangle$ is not a fuzzy sequential pattern, since $z1 \leq z2$ and there exist $j1 = 1$ and $j2 = 2$ such that $L_{b,j1} \subseteq L_{a,1}$ (i.e., $A^{\text{Product 1}}_{2,1} \subseteq A^{\text{Product 1}}_{2,1}$) and $L_{b,j2} \subseteq L_{a,2}$ (i.e., $A^{\text{Product 1}}_{2,2} \times A^{\text{Product 2}}_{2,1} \subseteq A^{\text{Product 1}}_{2,2}$). That is, $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2}\rangle$ is contained in $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2} \times A^{\text{Product 2}}_{2,1}, A^{\text{Product 1}}_{2,1} \times A^{\text{Product 2}}_{2,2}\rangle$. There is no doubt that the information contained in the latter sequence (i.e., $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2} \times A^{\text{Product 2}}_{2,1}, A^{\text{Product 1}}_{2,1} \times A^{\text{Product 2}}_{2,2}\rangle$) is more than that contained in the former sequence (i.e., $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2}\rangle$). We note that $\langle A^{\text{Product 1}}_{2,1}, A^{\text{Product 1}}_{2,2} \times A^{\text{Product 2}}_{2,1}, A^{\text{Product 1}}_{2,1} \times A^{\text{Product 2}}_{2,2}\rangle$ is a fuzzy sequential pattern if it is not contained in the other frequent fuzzy sequences.

**Example 9.** In this example, we demonstrate the possible application of fuzzy sequential pattern. We assume that $\langle A^{\text{Product 2}}_{2,2}, A^{\text{Product 1}}_{2,1}\rangle$ is a fuzzy sequential pattern. If product 1 is orange juices, and product 2 is apple juices, then a piece of useful information extracted from this pattern demonstrates that small purchase amounts of orange juices are likely to be bought by customers next time after they bought large purchase amounts of apple juices. This information can help decision makers (e.g., retailers) plan marketing strategy. For example, those customers, who have ever bought large purchase amounts of apple juices, may be attracted to buy more orange juice on sale.

## 4. Implementation of the Proposed Method

Based on the framework illustrated in Section 3, we present the real implementation of the proposed method in detail. As we have mentioned above, the first phase is to find frequent fuzzy sequences, and the second phase is to discover fuzzy sequential patterns by those frequent fuzzy sequences. Phases I and II are described in Sections 4.1 and 4.2, respectively.

### *4.1 Phase I: Frequent Fuzzy 1-Sequence Mining*

Table FGTTFS is used to generate frequent fuzzy grids, and consists of the following substructures:

(a) Fuzzy Grids table (FG): FG is a two-valued matrix. In FG, each row represents a fuzzy grid, while each column represents a linguistic value $A^{x_m}_{K,i_m}$.
(b) Transaction table (TT): each column represents $t^{(r)}_p$, while each element records the membership degree which $t^{(r)}_p$ belongs to the corresponding fuzzy grid.
(c) Column FS: stores the fuzzy support corresponding to the fuzzy grid in FG.

An initial tabular FGTTFS is shown in Table 1 as an example. From Table 1, we can see that there are two transaction records, $t^{(1)}_1$ and $t^{(1)}_2$ (i.e., $\alpha_1 = 2$), corresponding to customer 1, and two quantitative attributes $x_1$ and $x_2$ in a given database relation. Each attribute is partitioned into 2 partitions (i.e., $K = 2$). We can see that each element of FG is assigned to 0 or 1. Thus, we can apply Boolean operations on FG[$u$] = (FG[$u$,1], FG[$u$,2], FG[$u$,3], FG[$u$,4]) and FG[$v$] = (FG[$v$,1], FG[$v$,2], FG[$v$,3], FG[$v$,4]) (i.e., $u$-th row and $v$-th row of FG). For example, if we apply the *OR* operation on two rows, say FG[1] = (1, 0, 0, 0) and FG[3] = (0, 0, 1, 0), then (FG[1] *OR* FG[3]) = (1, 0, 1, 0) corresponding to a candidate 2-dim fuzzy grid $A^{x_1}_{2,1} \times A^{x_2}_{2,1}$, is generated. Then, $\text{FS}(A^{x_1}_{2,1} \times A^{x_2}_{2,1}) = \text{TT}[1] \cdot \text{TT}[3] = \max\{\mu^{x_1}_{2,1}(t^{(1)}_{1_1}) \cdot \mu^{x_2}_{2,1}(t^{(1)}_{1_2}), \mu^{x_1}_{2,1}(t^{(1)}_{2_1}) \cdot \mu^{x_2}_{2,1}(t^{(1)}_{2_2})\}$ is obtained to compare with the min FS. If $A^{x_1}_{2,1} \times A^{x_2}_{2,1}$ is frequent, then the data (i.e. FG[1] *OR* FG[3], $\text{TT}_3[1] \cdot \text{TT}_3[3]$, and $\text{FS}(A^{x_1}_{2,1} \times A^{x_2}_{2,1})$) will be inserted to corresponding data structures (i.e., FG, TT, and FS).

In the Apriori algorithm, two frequent $(l - 1)$-itemsets are joined to be a candidate l-itemset ($3 \leq l \leq$ d), and these two frequent itemsets share $(l - 2)$ items Agrawal et al., 1993; Jang & Sun, 1995). Similarly, a candidate $l$-dim fuzzy grid, say $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-1}}_{K,i_{l-1}} \times A^{x_l}_{K,i_l}$, is also derived by joining two frequent ($l$–1)-dim fuzzy grids (e.g., $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-2}}_{K,i_{l-2}} \times A^{x_l}_{K,i_l}$ and $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-2}}_{K,i_{l-2}} \times A^{x_{l-1}}_{K,i_{l-1}}$), and these two frequent grids share a frequent ($l$ − 2)-dim fuzzy grid (i.e., $A^{x_1}_{K,i_1} \times A^{x_2}_{K,i_2} \times \cdots \times A^{x_{l-2}}_{K,i_{l-2}}$). For example, we can use two frequent fuzzy grids $A^{x_1}_{3,2} \times A^{x_2}_{3,1}$ and $A^{x_1}_{3,2} \times A^{x_3}_{3,3}$ to generate the candidate 3-dim fuzzy grid $A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ because $A^{x_1}_{3,2} \times A^{x_2}_{3,1}$ and $A^{x_1}_{3,2} \times A^{x_3}_{3,3}$ share the same 1-dim fuzzy grid $A^{x_1}_{3,2}$.

However, $A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ can also be generated by joining $A^{x_1}_{3,2} \times A^{x_2}_{3,1}$ with $A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ since $A^{x_1}_{3,2} \times A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ is a subset of $A^{x_2}_{3,1} \times A^{x_3}_{3,3}$, and $A^{x_1}_{3,2} \times A^{x_2}_{3,1}$ and $A^{x_2}_{3,1} \times A^{x_3}_{3,3}$ share $A^{x_2}_{3,1}$. This means that we must select one of many possible combinations to avoid redundant computations. To resolve this problem, we consider that if there exist $l$ integers number $e_1, e_2, \cdots, e_{l-1}, e_l$ where $1 \le e_1 < e_2 < \cdots < e_{l-1} < e_l \le d$, such that FG[$u$, $e_1$] = FG[$u$, $e_2$] = $\cdots$ = FG[$u$, $e_{l-2}$] = FG[$u$, $e_{l-1}$] = 1 and FG[$v$, $e_1$] = FG[$v$, $e_2$] = $\cdots$ = FG[$v$, $e_{l-2}$] = FG[$v$, $e_l$] = 1, where FG[$u$] and FG[$v$] correspond to individual frequent ($l$ − 1)-dim fuzzy grids, then FG[$u$] and FG[$v$] can be merged to generate a candidate $l$-dim fuzzy grid. It should be noted that any two 1-dim fuzzy grids defined in the same attribute cannot be simultaneously contained in a candidate $l$-dim fuzzy grid ($l \ge 2$). Thus (1, 1, 0, 0) or (0, 0, 1, 1) are invalid. We describe the detailed procedure of phase I as follows.

**Algorithm:** The proposed method for discovering fuzzy sequential patterns (phase I)
**Input:** a. A specified database;
b. Minimum fuzzy support.
**Output:** Frequent fuzzy 1-sequences (i.e., frequent fuzzy grids)
**Method:**
Step 1. Perform the simple fuzzy partition method
Step 2. Scan the database and then construct the initial FGTTFS
Step 3. Generate frequent fuzzy grids
3-1. Generate frequent 1-dim fuzzy grids
Set $l = 1$ and eliminate the rows of initial FGTTFS corresponding to candidate 1-dim fuzzy grids that are not frequent.
3-2. Generate frequent $l$-dim fuzzy grids
Set $l + 1$ to $l$. If there is only one frequent ($l$ − 1)-dim fuzzy grid, then go to phase II.
*For two unpaired rows, FGTTFS[u] and FGTTFS[v] (u ≠ v), corresponding to frequent (l − 1)-dim fuzzy grids do*
Compute (FG[$u$] OR FG[$v$]) corresponding to a candidate $l$-dim fuzzy grid $c$.
3-2-1. Examine the validity of $c$
If any two 1-dim fuzzy grid defined in the same attribute, then discard $c$ and skip Steps 3-2-2, 3-2-3 and 3-2-4. That is, $c$ is invalid.
3-2-2. If any two 1-dim fuzzy grids defined in the same attribute for $c$, then discard $c$ and skip Steps 3-2-3 and 3-2-4.

TABLE 2. Table FSEFS for an example.

| Fuzzy sequence | FSE | | | | FS |
|---|---|---|---|---|---|
| | $A^{x_1}_{2,1}$ | $A^{x_1}_{2,2}$ | $A^{x_2}_{2,1}$ | $A^{x_1}_{2,1} \times A^{x_2}_{2,1}$ | |
| $\langle A^{x_1}_{2,1}\rangle$ | 1 | 0 | 0 | 0 | FS($\langle A^{x_1}_{2,1}\rangle$) |
| $\langle A^{x_1}_{2,2}\rangle$ | 0 | 1 | 0 | 0 | FS($\langle A^{x_1}_{2,2}\rangle$) |
| $\langle A^{x_2}_{2,1}\rangle$ | 0 | 0 | 1 | 0 | FS($\langle A^{x_2}_{2,1}\rangle$) |
| $\langle A^{x_1}_{2,1} \times A^{x_2}_{2,1}\rangle$ | 0 | 0 | 0 | 1 | FS($\langle A^{x_1}_{2,1} \times A^{x_2}_{2,1}\rangle$) |

3-2-3. Examine whether or not FG[$u$] and FG[$v$] share ($l$ − 2) 1's
If there exist integers $1 \le e_1 < e_2 < \cdots < e_{l-1} < e_l \le d$ such that FG[$u$, $e_1$] = FG[$u$, $e_2$] = $\cdots$ = FG[$u$, $e_{l-2}$] = FG[$u$, $e_{l-1}$] = 1 and FG[$v$, $e_1$] = FG[$v$, $e_2$] = $\cdots$ = FG[$v$, $e_{l-2}$] = FG[$_{v,\ e}$l] = 1, then compute (TT[$e_1$] ( TT[$e_2$] $\cdot \cdots \cdot$ TT[$e_l$]) and fuzzy support denoted by *fs* of $c$.
3-2-4. Examine the fuzzy support of the newly generated candidate fuzzy grid Insert (FG[$u$] OR FG[$v$]) to FG, (TT[$e_1$] ( TT[$e_2$] $\cdot \cdots \cdot$ TT[$e_l$]) to TT and *fs* to FS when *fs* is larger than or equal to min FS; otherwise, discard $c$.
*End*
3-3. Check whether any frequent $l$-dim fuzzy grid is generated or not
If any frequent $l$-dim fuzzy grid is generated, then go to Step 3-2; otherwise go to phase II. It is noted that the final FGTTFS stores only frequent fuzzy grids.

### 4.2 Phase II: Fuzzy Sequential Pattern Mining

Table FSEFS is used to generate fuzzy sequences and consists of two substructures including the fuzzy sequences table (FSE) and the column FS. FSE is an integer matrix and each row represents a fuzzy sequence, while each column represents a frequent fuzzy grid. FSE can allow us to easily determine which fuzzy sequence is generated and which frequent fuzzy grids are contained in this sequence.

We assume that the initial FSEFS is generated as Table 2, where we can see that four frequent fuzzy 1-sequences are generated. By using the asymmetric aggregation operator $\oplus$ for FSE[$u$] and FSE[$v$] (i.e., FSE[$u$] $\oplus$ FSE[$v$] or FSE[$v$] $\oplus$ FSE[u]), we can obtain a candidate fuzzy sequence. FSE[$u$] $\oplus$ FSE[$v$] is computed as follows:

$$\mathrm{FSE}[u,j] \oplus \mathrm{FSE}[v,j] = \left\{ \begin{array}{ll} \mathrm{FSE}[u,j]+1, & \text{if } 0 \le \mathrm{FSE}[v,j] \le \mathrm{FSE}[u,j], \\ & \quad \mathrm{FSE}[u,j] \ne 0 \\ 1, & \text{if } \mathrm{FSE}[u,j] < \mathrm{FSE}[v,j] \\ 0, & \text{if } 0 = \mathrm{FSE}[u,j] = \mathrm{FSE}[v,j] \end{array} \right\} \quad (12)$$

**Example 10.** (2, 1, 0, 0) corresponding to a candidate fuzzy 2-sequence $\langle A^{x_1}_{2,1}, A^{x_1}_{2,2}\rangle$ is generated by FSE[1] $\oplus$ FSE[2] (i.e. (1, 0, 0, 0) $\oplus$ (0, 1, 0, 0)) in Table 2, since

FSE[2, 1] ≤ FSE[1, 1], FSE[1, 2] < FSE[2, 2], FSE[1, 3] = FSE[2, 3] = FSE[1, 4] = FSE[2, 4] = 0. The frequent fuzzy grid corresponding to the largest number (i.e., 2) in FSE[1] $\oplus$ FSE[2] is the first item (i.e. first occurrence) of this sequence, the frequent fuzzy grid corresponding to the next-to-the-largest number (i.e., 1) is the second item (i.e., second occurrence) of this sequence, and so on. If $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$ is frequent, then corresponding data (i.e., FSE[1] $\oplus$ FSE[2] and FS($\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$) will be inserted to corresponding data structures (i.e., FSE and FS).

A candidate fuzzy $k$-sequence ($k \geq 3$) can be derived by joining two frequent fuzzy ($k - 1$)-sequences, and these two sequences share a frequent fuzzy ($k - 2$)-sequence. For example, we can use two frequent fuzzy sequences, say $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$ and $\langle A_{2,1}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$, to generate a candidate fuzzy 3-sequence $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$ because $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$ and $\langle A_{2,1}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$ share the same fuzzy 1-sequence $\langle A_{2,1}^{x_1}\rangle$. Actually, another candidate fuzzy 3-sequence $\langle A_{2,1}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}, A_{2,2}^{x_1}\rangle$ can be generated.

However, $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$ can also be generated by combining $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$ with $\langle A_{2,2}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$ since $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$ and $\langle A_{2,2}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$ share $\langle A_{2,2}^{x_1}\rangle$. To resolve this problem, we consider that that if there exist ($k - 2$) (k ≥ 3) integers numbers $e_1, e_2, \ldots, e_{k-3}, e_{k-2}$ where $1 \leq e_1, e_2, \ldots, e_{k-2} \leq \beta$, such that FSE[$u$, $e_1$] = FSE[$v$, $e_1$] = 2, FSE[$u$, $e_2$] = FSE[$v$, $e_2$] = 3, . . . , FSE[$u$, $e_{k-3}$] = FSE[$v$, $e_{k-3}$] = $k - 2$ and FSE[$u$, $e_{k-2}$] = FSE[$v$, $e_{k-2}$] = $k - 1$, where FSE[$u$] and FSE[$v$] correspond to individual frequent fuzzy ($k - 1$)-sequences, then we can employ (FSE[$u$] $\oplus$ FSE[$v$]) and (FSE[$v$] $\oplus$ FSE[$u$]) to generate various candidate fuzzy $k$-sequences.

**Example 11.** FSE[$u$] $\oplus$ FSE[$v$] = (1, 0, 0, 2) (i.e., $\langle A_{2,1}^{x_1} \times A_{2,1}^{x_2}, A_{2,1}^{x_1}\rangle$) $\oplus$ (0, 1, 0, 2) (i.e., $\langle A_{2,1}^{x_1} \times A_{2,1}^{x_2}, A_{2,2}^{x_1}\rangle$) = (2, 1, 0, 3) (i.e., $\langle A_{2,1}^{x_1} \times A_{2,1}^{x_2}, A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$) can be obtained since there exist (3 − 2) (i.e., $k$ = 3) integers numbers $e_1$ = 3 such that FG[$u$, 4] = FG[$v$, 4] = 2. As we mentioned above, FSE[$v$] $\oplus$ FSE[$u$] = (0, 1, 0, 2) $\oplus$ (1, 0, 0, 2) = (1, 2, 0, 3) (i.e., $\langle A_{2,1}^{x_1} \times A_{2,1}^{x_2}, A_{2,2}^{x_1}, A_{2,1}^{x_1}\rangle$) can be also obtained. However, for FSE[$u$] = (2, 1, 0, 0) (i.e., $\langle A_{2,1}^{x_1}, A_{2,2}^{x_1}\rangle$) and FSE[$v$] = (0 , 2, 0, 1) (i.e., $\langle A_{2,2}^{x_1}, A_{2,1}^{x_1} \times A_{2,1}^{x_2}\rangle$), (2, 1, 0, 0) $\oplus$ (0 , 2, 0, 1) is invalid since FSE[$u$, $j$] ≠ FSE[$v$, $j$] ($j$ = 1, 2, 3, 4).

Additionally, we can observe that if a candidate fuzzy $k$-sequence is generated by FSE[$u$] $\oplus$ FSE[$v$], then FSE[$u$] $\oplus$ FSE[$v$] must contain $k$ positive and consecutive positive integer numbers (i.e., 1, 2, . . . , $k - 1$, $k$). For example, (1, 2, 3, 0) contain 3 various integer numbers (i.e., 1, 2, 3). We describe the detailed procedure of phase II as follows.

**Algorithm:** The proposed method for discovering fuzzy sequential patterns (phase II)
**Input:** a. Frequent fuzzy 1-sequences (initial FSEFS)
b. Minimum fuzzy support.
**Output:** Fuzzy sequential patterns
**Method:**
**Step 0. Initialization**
Set 1 to $k$.

**Step 1. Generate frequent fuzzy $k$-sequences ($2 \leq k \leq \beta$)**
Set $k + 1$ to $k$. If there is only one frequent fuzzy ($k - 1$)-sequence, then go to Step 2. *For two unpaired rows, FSEFS[u] and FSEFS[v] (u &neq; v), corresponding to frequent fuzzy (k − 1)- sequences do*

1-1. Examine the validity of FSE[u] ( FSE[v] If there exist ($k - 2$) integers numbers $e_1, e_2, \ldots, e_{k-3}, e_{k-2}$ where $1 \leq e_1, e_2, \ldots, e_{k-2} \leq \beta$, such that FSE[$u$, $e_1$] = FSE[$v$, $e_1$] = 2, FSE[$u$, $e_2$] = FSE[$v$, $e_2$] = 3, . . . , FSE[$u$, $e_{k-3}$] = FSE[$v$, $e_{k-3}$] = $k - 2$ and FSE[$u$, $e_{k-2}$] = FSE[$v$, $e_{k-2}$] = $k - 1$, then compute FSE[$u$] $\oplus$ FSE[$v$] and FSE[$v$] $\oplus$ FSE[$u$] to generate candidate fuzzy $k$-sequences s′ and s″, respectively

1-2. Examine the fuzzy support of the newly generated fuzzy sequences
Insert FSE[$u$] $\oplus$ FSE[$v$] or FSE[$v$] $\oplus$ FSE[$u$] to FSE, and FS(s′) or FS(s″) to FS when FS(s′) or FS(s″) is larger than or equal to min FS.

*End*

**Step 2. Check whether or not any frequent fuzzy $k$-sequence is generated**
If any frequent fuzzy $k$-sequence is generated, then repeat by going to Step 1, else go to Step 3.

**Step 3. Find fuzzy sequential patterns**
For any two rows of FSEFS, say FSE[$u$] and FSE[$u$], if the fuzzy sequence corresponding to FSE[$u$] is contained in the fuzzy sequence corresponding to FSE[$v$] by using the method introduced in Subsection 3.3, then eliminate FSE[$u$]. It is noted that the final FSEFS stores only fuzzy sequential patterns.

The analysis of the computational complexity is somewhat difficult. We roughly analyze the proposed method in the worst case (i.e., minimum fuzzy support = 0) since an appropriate minimum fuzzy support is determined by decision makers. As we have mentioned above, the smaller minimum fuzzy support is specified by users, the larger number of frequent fuzzy 1-sequences will be generated. It is thus clear that the more frequent fuzzy 1-sequences are generated, the more computational steps will be used in phase I. For example, in Figure 4, there are 15 possible frequent fuzzy grids (i.e., 6 1-dim grids and 9 2-dim grids) in the worst case. It is also clear that 15 is larger than $3^1 + 3^2$. In a similar manner, if each of the $d$ quantitative attributes is partitioned into $K$ linguistic values, then the total number of frequent fuzzy grids would be larger than $K^1 + K^2 + \cdots + K^d$ in the worse case. That is, it is possible that the number of processing computational steps in phase I will increase exponentially for databases with high degree. If let $\beta$ be equal to $K^1 + K^2 + \cdots + K^d$, then $C_2^{\beta}$ iterations of the for-loop in Step 1 of phase II would be performed.

In the following section, two numerical examples are utilized to demonstrate the usefulness of the proposed method.

TABLE 3. Table BOUGHT sorted by transaction time for each customer.

| Record | Transaction time | Product 1 | Product 2 | Product 3 | Product 4 | Product 5 | Product 6 | Product 7 | Product 8 | Product 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $t_1^{(1)}$ | 04/10/02 | * | * | 5 | * | * | * | * | * | * |
| $t_2^{(1)}$ | 05/11/02 | * | * | * | * | * | * | * | * | 8 |
| $t_1^{(2)}$ | 04/12/02 | 6 | 10 | 9 | * | * | * | * | * | * |
| $t_2^{(2)}$ | 04/25/02 | * | * | 8 | * | * | * | 14 | * | * |
| $t_3^{(2)}$ | 06/01/02 | * | * | 1 | 9 | * | 12 | 6 | * | 6 |
| $t_1^{(3)}$ | 05/02/02 | * | 8 | 7 | * | 10 | * | 9 | * | * |
| $t_1^{(4)}$ | 04/05/02 | * | * | 15 | * | * | * | 12 | * | * |
| $t_2^{(4)}$ | 04/29/02 | * | 6 | * | 10 | * | * | 10 | * | * |
| $t_3^{(4)}$ | 06/02/02 | * | * | 4 | * | * | * | * | * | 12 |
| $t_1^{(5)}$ | 05/20/02 | * | * | * | * | * | * | * | * | 5 |

## 5. Numerical Examples

The main purpose of this section is to show the usefulness of the proposed method. Two possible applications relating to analyze purchase behaviors are demonstrated as follows: one to analyze general purchase behaviors, and the other to analyze purchase behaviors of one group.

### *A. Analysis of General Purchase Behaviors*

In a supermarket or a mart, a table named BOUGHT with 10 transactions is extracted from transaction databases as shown in Table 3, where the asterisk denotes that one product was not purchased in that transaction. We can see that $\alpha_1 = 2$, $\alpha_2 = 3$, $\alpha_3 = 1$, $\alpha_4 = 3$, $\alpha_5 = 1$.

There are nine quantitative attributes, and each quantitative attribute ranging from zero to twenty is partitioned into three linguistic values (i.e., $K = 3$), which are similar to partitions depicted in Figure 2, by the simple fuzzy partition method. Therefore, fuzzy subsets defined in individual partitions can be linguistically interpreted, such as for the product $m$ ($m = 1 \cdots 9$):

$$A_{3,1}^{\text{Product } m} : \text{small}$$

$$A_{3,2}^{\text{Product } m} : \text{medium}$$

$$A_{3,3}^{\text{Product } m} : \text{large}$$

Then, we employ the proposed method to find fuzzy sequential patterns from BOUGHT by specifying min FS to be 0.20. The detailed computation process is omitted for simplicity. At the end of phase II, 10 fuzzy sequential patterns with individual fuzzy supports can be discovered (i.e., $\langle A_{3,2}^{\text{Product 5}} \rangle$ with 0.20, $\langle A_{3,1}^{\text{Product 9}} \rangle$ with 0.20, $\langle A_{3,2}^{\text{Product 2}} \times A_{3,2}^{\text{Product 3}} \rangle$ with 0.29, $\langle A_{3,1}^{\text{Product 3}} \times A_{3,2}^{\text{Product 9}} \rangle$ with 0.20, $A_{3,2}^{\text{Product 3}} \times A_{3,2}^{\text{Product 7}} \rangle$ with 0.30, $\langle A_{3,2}^{\text{Product 2}}, A_{3,2}^{\text{Product 3}} \rangle$ with 0.21, $\langle A_{3,2}^{\text{Product 3}}, A_{3,1}^{\text{Product 3}} \rangle$ with 0.22, $\langle A_{3,2}^{\text{Product 7}}, A_{3,1}^{\text{Product 3}} \rangle$ with 0.20, $\langle A_{3,2}^{\text{Product 7}}, A_{3,2}^{\text{Product 9}} \rangle$ with 0.20, and $\langle A_{3,3}^{\text{Product 7}}, A_{3,2}^{\text{Product 9}} \rangle$ with 0.21). It is suggested that decision makers should pay more attention on those patterns with larger fuzzy support.

The aforementioned patterns can help decision makers plan marketing strategies. For example, $\langle A_{3,2}^{\text{Product 3}}, A_{3,1}^{\text{Product 3}} \rangle$ and $\langle A_{3,2}^{\text{Product 7}}, A_{3,1}^{\text{Product 3}} \rangle$ demonstrate that small purchase amounts of product 3 are likely to be bought by customers next time after they bought medium purchase amounts of product 3 or medium purchase amounts of product 7. Those customers, which can be found from databases by a query language such as SQL, may be attracted to buy more product 3 on sale. Decision makers should further analyze the possible reasons why small purchase amounts of product 3 are likely to be bought.

Another interesting patterns are $\langle A_{3,2}^{\text{Product 2}}, A_{3,2}^{\text{Product 3}} \rangle$ and $\langle A_{3,2}^{\text{Product 3}}, A_{3,1}^{\text{Product 3}} \rangle$. That is, we find that $\langle A_{3,2}^{\text{Product 2}}, A_{3,2}^{\text{Product 3}}, A_{3,1}^{\text{Product 3}} \rangle$ is not generated. Those customers who bought medium purchase amounts of product 2 may feel that medium purchase amounts of product 3 can sufficiently satisfy their requirement. Maybe, they can be also attracted to buy more product 3 on sale. It is possible that $A_{3,3}^{\text{Product 3}}$ will be generated in the next pattern mining by performing an appropriate marketing strategy.

### *B. Analysis of Purchase Behaviors of One Group*

It is possible that one group is very significant for a business. The analysis of purchase behaviors of one group is thus meaningful. If we treat Table 3 as transaction records of one group, say a group of high salary (GHS), then a new table sorted by transaction time for GHS is generated as shown in Table 4. It is also reasonable that we view GHS as customer 1 such that $\alpha_1 = 10$. Each quantitative attribute ranging from zero to twenty is still partitioned into three linguistic values. Then, we employ the proposed method to find fuzzy sequential patterns from BOUGHT by specifying min FS to be 0.90.

At the end of phase II, 4 fuzzy sequential patterns with individual fuzzy supports can be found (i.e., $\langle A_{3,2}^{\text{Product 2}} \times A_{3,2}^{\text{Product 3}}, A_{3,2}^{\text{Product 5}} \rangle$ with 0.90, $\langle A_{3,2}^{\text{Product 2}} \times A_{3,2}^{\text{Product 3}}, A_{3,3}^{\text{Product 7}} \rangle$ with 0.90, $\langle A_{3,2}^{\text{Product 2}}, A_{3,3}^{\text{Product 7}}, A_{3,2}^{\text{Product 5}} \times A_{3,2}^{\text{Product 7}} \rangle$ with 0.90, and $\langle A_{3,2}^{\text{Product 2}}, A_{3,3}^{\text{Product 7}}, A_{3,2}^{\text{Product 5}}, A_{3,1}^{\text{Product 3}} \rangle$ with 0.90). Decision-makers should make use of these patterns to plan appropriate marketing strategies for GHS. The possible analysis of fuzzy sequential patterns is omitted here.

TABLE 4. Table BOUGHT sorted by transaction time for one group.

| Transaction time | Product 1 | Product 2 | Product 3 | Product 4 | Product 5 | Product 6 | Product 7 | Product 8 | Product 9 |
|---|---|---|---|---|---|---|---|---|---|
| 04/05/02 | * | * | 15 | * | * | * | 12 | * | * |
| 04/10/02 | * | * | 5 | * | * | * | * | * | * |
| 04/12/02 | 6 | 10 | 9 | * | * | * | * | * | * |
| 04/25/02 | * | * | 8 | * | * | * | 14 | * | * |
| 04/29/02 | * | 6 | * | 10 | * | * | 10 | * | * |
| 05/02/02 | * | 8 | 7 | * | 10 | * | 9 | * | * |
| 05/11/02 | * | * | * | * | * | * | * | * | 8 |
| 05/20/02 | * | * | * | * | * | * | * | * | 5 |
| 06/01/02 | * | * | 1 | 9 | * | 12 | 6 | * | 6 |
| 06/02/02 | * | * | 4 | * | * | * | * | * | 12 |

## 6. Discussions and Conclusions

As we have mentioned above, the main aim of this paper is to propose a fuzzy data mining technique to discover fuzzy sequential patterns by using the simple partition method. The first phase is to find purchase behaviors that frequently occurred for a period of time (i.e., frequent fuzzy grids), and the second phase is to discover fuzzy sequential patterns by analyzing the temporal relation among those purchase behaviors found in the first phase. Two numerical examples can demonstrate the usefulness and possible applications of the proposed method. Several improvements or suggestions of the proposed method are discussed as follows. Some issues are left for future works.

The rough analysis of the computational complexity in the worst case is described in Section 4. Since it is possible that the number of processing computational steps in phase I will increase exponentially for databases with high degree, the computational time and the storage requirement will be enlarged. The usefulness of the proposed method for databases with high degree is more or less deteriorated. Therefore, it is possible to remove the unimportant attributes to reduce the dimensions. In reducing the feature space dimensions, several feature selection methods can be used such as the principal component analysis (Sharma, 1996), which is a useful multivariate analysis technique. Additionally, various linguistic interpretations of a fuzzy set can be obtained by linguistic hedges (Zimmermann, 1991, 1996; Ishibuchi & Nii, 2001; Pedrycz & Gomide, 1998) for $A^{x_m}_{K,i_m}$ such as "very $A^{x_m}_{K,i_m}$" denoted by $(A^{x_m}_{K,i_m})'$ or "more or less $A^{x_m}_{K,i_m}$" denoted by $(A^{x_m}_{K,i_m})''$ as follows:

$$(A^{x_m}_{K,i_m})' = \text{very} \quad A^{x_m}_{K,i_m} = (A^{x_m}_{K,i_m})^2 \tag{13}$$

$$(A^{x_m}_{K,i_m})'' = \text{more or less} \quad A^{x_m}_{K,i_m} = (A^{x_m}_{K,i_m})^{1/2} \tag{14}$$

The membership functions of $(A^{x_m}_{K,i_m})'$ and $(A^{x_m}_{K,i_m})''$ can be stated as $[\mu^{x_m}_{K,i_m}(x)]^2$ and $[\mu^{x_m}_{K,i_m}(x)]^{1/2}$, respectively. Therefore, there are three different linguistic terms defined in each partition, such as $A^{x_m}_{K,i_m}$, "very $A^{x_m}_{K,i_m}$" and "more or less $A^{x_m}_{K,i_m}$." Therefore, there are $3K$ different linguistic values defined in $x_m$. It is possible that the proposed method uses these linguistic values simultaneously to discover fuzzy sequential patterns. We believe that such extensions will make fuzzy sequential patterns to be more versatile and more useful for users.

As we have mentioned in Section 2, the decision makers can subjectively determine the number, locations and shapes of fuzzy sets in each quantitative attribute depending on their preferences, past experiences, or prior knowledge. The advantage is that the fuzzy sequential patterns are more comprehensible for the decision-makers. That is, it is not necessary to provide methods to find general or optimal parameter specifications (i.e., number, locations, and shapes) of membership functions in each quantitative attribute.

## Acknowledgements

We are very grateful to the anonymous referees for their valuable comments and constructive suggestions.

## References

Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in frequent databases. In Proceedings of the ACM SIGMOD International Conference on Management of Data, 1993, (pp. 207-216). Washington D.C.

Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., & Verkamo, A.I. (1996). Fast discovery of association rules. In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.). Advances in knowledge discovery and data mining, Menlo Park: AAAI Press, 1996, pp. 307–328.

Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In Proceedings of the Eleventh International Conference on Data Engineering, 1995, (pp. 3–14). Taipei, Taiwan.

Berry, M., & Linoff, G. (1997). Data mining techniques: for marketing, sales, and customer support. New York: John Wiley & Sons.

Bezdek J.C. (1981). Pattern recognition with fuzzy objective function algorithms, New York: Plenum.

Chen, S.M., & Jong, W.T. (1997). Fuzzy query translation for relational database systems. IEEE Transactions on Systems, Man, and Cybernetics, 27(4), 714–721.

Han J.W., & Kamber M. (2001). Data mining: concepts and techniques. San Francisco: Morgan Kaufmann.

Homaifar, A., & McCormick, E. (1995). Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. IEEE Transactions on Fuzzy Systems, 3(2), 129–139.

Hong T.P., Wang T.T., Wang, S.L., & Chien, B.C. (2000). Learning a coverage set of maximally general fuzzy rules by rough sets. Expert Systems with Applications, 19(2), 97–103.

Hu, Y.C., Chen, R.S., & Tzeng, G.H. Generating learning sequences for decision makers through data mining and competence set expansion. IEEE Transactions on Systems, Man, and Cybernetics, Part B, 32(5), 679–686.

Hu Y.C., Tzeng G.H., & Chen R. S. (2002). Mining fuzzy association rules for classification problems. Computers and Industrial Engineering, 43(4), 735–750.

Ishibuchi, H., Nakashima, T., Murata, T. (1999). Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. IEEE Transactions on Systems, Man, and Cybernetics, 29(5), 601–618.

Ishibuchi, H., Nakashima T., & Murata, T. (2001). Three-objective genetics-based machine learning for linguistic rule extraction. Information Sciences, 136, 109–133.

Ishibuchi, H., Nozaki K., Yamamoto N., & Tanaka H. (1995). Selecting fuzzy if-then rules for classification problems using genetic algorithms. IEEE Transactions on Fuzzy Systems, 3(3), 260–270.

Ishibuchi H., & Nii, M. (2001). Numerical analysis of the learning of fuzzified neural networks from fuzzy if-then rules. Fuzzy Sets and Systems 120, 281–307.

Ishibuchi, H., Yamamoto, T., & Nakashima, T. (2001). Fuzzy data mining: effect of fuzzy discretization. In Proceedings of the 1st IEEE International Conference on Data Mining, San Jose, CA, 2001, pp.241–248.

Jang, J.S.R. (1993). ANFIS: adaptive-network-based fuzzy inference systems. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665–685.

Jang, J.S.R., & Sun, C.T. (1995). Neuro-fuzzy modeling and control. Proceedings of the IEEE, 83(3), 378–406.

Myra, S. (2000). Web usage mining for web site evaluation. Communications of the ACM, 43(8), 127–134.

Pedrycz, W. (1994). Why triangular membership functions? Fuzzy Sets and Systems 64, 21–30.

Pedrycz, W., & Gomide, F. (1998). An Introduction to Fuzzy Sets: Analysis and Design, Cambridge: MIT Press.

Ravi, V., Zimmermann, H. -J. (2000). Fuzzy rule based classification with *FeatureSelector* and modified threshold accepting. European Journal of Operational Research, 123(1), 16–28.

Saaty, T.L. (1980). The analytic hierarchy process: planning, priority setting, resource allocation. New York: McGraw-Hill.

Sharma, S. (1996). Applied multivariate techniques. Singapore: John Wiley & Sons.

Sun, C.T. (1994). Rule-base structure identification in an adaptive-network-based fuzzy inference system. IEEE Transactions on Fuzzy Systems, 2(1), 64–73.

Zadeh, L.A. (1965). Fuzzy sets. Information Control, 8(3), 338–353.

Zadeh, L.A. (1975, 1976) The concept of a linguistic variable and its application to approximate reasoning, Information Science, 8(3), 199–249(part 1); 8(4) 301–357(part 2); 9(1), 43–80(part 3).

Zimmermann, H.-J. (1991). Fuzzy sets, decision making, and expert systems. Boston: Kluwer.

Zimmermann, H.-J. (1996). Fuzzy set theory and its applications. Boston: Kluwer.

Wang, L.X., & Mendel, J.M. (1992). Generating fuzzy rules by learning from examples. IEEE Transactions on Systems, Man, and Cybernetics, 22(6), 1414–1427.