amcs

EVOLUTIONARY ALGORITHMS AND FUZZY SETS FOR DISCOVERING TEMPORAL RULES

STEPHEN G. MATTHEWS *, MARIO A. GONGORA **, ADRIAN A. HOPGOOD ***

* Intelligent Systems Laboratory, Department of Engineering Mathematics
University of Bristol, Bristol, BS8 1UB, UK
e-mail: `stephen.matthews@bristol.ac.uk`

**Centre for Computational Intelligence
De Montfort University, The Gateway, Leicester, LE1 9BH, UK
e-mail: `mgongora@dmu.ac.uk`

***Sheffield Business School
Sheffield Hallam University, Sheffield, S1 1WB, UK
e-mail: `a.hopgood@shu.ac.uk`

A novel method is presented for mining fuzzy association rules that have a temporal pattern. Our proposed method contributes towards discovering temporal patterns that could otherwise be lost from defining the membership functions before the mining process. The novelty of this research lies in exploring the composition of fuzzy and temporal association rules, and using a multi-objective evolutionary algorithm combined with iterative rule learning to mine many rules. Temporal patterns are augmented into a dataset to analyse the method's ability in a controlled experiment. It is shown that the method is capable of discovering temporal patterns, and the effect of Boolean itemset support on the efficacy of discovering temporal fuzzy association rules is presented.

Keywords: fuzzy association rules, temporal association rules, multi-objective evolutionary algorithm.

1. Introduction

Knowledge discovery in databases is the process of acquiring useful information from databases (Piatetsky-Shapiro, 1990). Data mining is one step of this process where the goal is to discover knowledge that is accurate, comprehensible and interesting (Freitas, 2002). There are two tasks of data mining that are distinguished by the use of the information discovered: predictive for classification or prediction tasks, and descriptive for providing information about patterns and relationships in data.

Association rule mining is a well-established method of data mining that identifies significant correlations between Boolean items in transactional data (Agrawal *et al*., 1993). Fuzzy sets (Zadeh, 1965) are used to model quantitative attributes with fuzzy association rule mining (Chan and Au, 1997). In the market basket problem, an example of a typical fuzzy association rule is shown below:

IF *pizza is high quantity* THEN *beer is high quantity*.

This paper uses fuzzy sets to represent numeric values to provide a smoother transition between boundaries of crisp intervals, and to better handle uncertainty. This is common in descriptive tasks such as association rule mining (e.g., Chan and Au, 1997; Hong *et al*., 2008), but also more generally in predictive tasks (e.g., Alcalá *et al*., 2007b; Casillas *et al*., 2005; Venturini, 1993).

For mining fuzzy association rules in the traditional manner, the following procedure is often used:

1. Define linguistic labels (Zadeh, 1975) and membership function parameters.

2. Mine the rules using the linguistic labels.

This two stage procedure presents an interesting problem because some temporal rules can be lost as a result of the first stage. The traditional method assumes

that the fuzzy association rules are static, meaning they do not change between when the first and last transactions occurred, and so they hold across the entire dataset. However, real-world datasets can have underlying temporal patterns.

There are temporal periods (e.g., a Friday) of the dataset where there can be more significant rules that use different membership functions. These different membership functions may not be captured in the traditional method because the measurement of support is taken across the entire dataset. These rules may occur at the intersection of membership function boundaries and so they are represented as less significant rules, despite being more prominent in a temporal period of the dataset. Less significant rules are lost because they fall below the minimum support threshold.

Figure 1 demonstrates how one item in a rule may have a membership function (marked with dashed line) that occurs more prominently in a temporal period of a dataset. This pattern is inadequately represented with the medium and high membership functions in that temporal period. Although traditional approaches do find temporal patterns of fuzzy association rules, they may not discover all significant patterns because some rules may contain membership functions that lie on the intersection of boundaries. For this reason, our proposed approach does not set the membership functions before mining to provide some flexibility for discovering these rules.
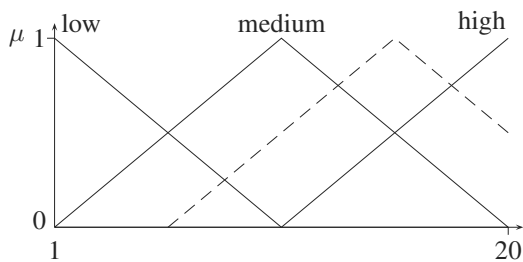

Fig. 1. Example of a membership function (dashed line) occurring at the intersection of two adjacent descriptive membership functions in a temporal period of a dataset.

This paper reports on a proposed method, and a methodology of analysis, for temporal fuzzy association rule mining that uses fuzzy numbers to overcome the problem of losing rules. The ability and efficacy of the following two aspects of the proposed approach are analysed:

- Discovery of a temporal fuzzy association rule in a *single* temporal period.

- Discovery of a temporal fuzzy association rule in *multiple* temporal periods.

Our proposed method outlines how to discover fuzzy association rules that can be lost with traditional

approaches. An approximate Fuzzy Rule-Based System (FRBS) approach is chosen because the flexibility in the definition of membership functions allows temporal fuzzy association rules to be discovered with fuzzy numbers. Simultaneously searching the itemset space, the temporal space and the membership function space is a complex problem, so a Genetic Algorithm (GA) is used. The search space is complex because of its size, and its continuous space (membership function parameters) and discrete space (market basket items). Searching this combination of search spaces with an exhaustive search method is prohibitive, because it requires partitioning the continuous space into fixed fuzzy sets. Our motive is to avoid fixing the fuzzy sets in order to provide flexibility in discovering rules that occur at the intersection of membership function boundaries.

A Multi-Objective Evolutionary Algorithm (MOEA) and Iterative Rule Learning (IRL) are used to mine multiple rules. Our previous work (Matthews *et al.*, 2011b; 2011a) is improved by changing the chromosome representation from itemsets to rules, modifying the fitness function, introducing a method of IRL for a multi-objective evolutionary algorithm, and an improved methodology for augmenting patterns.

This paper is organised as follows. Section 2 presents an overview of related works on association rule mining, Section 3 discusses the choice of an approximate FRBS, Section 4 describes the MOEA and IRL for mining temporal fuzzy association rules, Section 5 presents the experimental setup and methodology, Section 6 discusses results, and conclusions are drawn in Section 7.

2. Association rule mining

Association rule mining is an exploratory and descriptive rule induction process of identifying significant correlations between items in Boolean transaction datasets (Agrawal *et al.*, 1993) used for data analysis and interpretation. Association rules are expressed as an implication of the form $X \Rightarrow Y$ where the antecedent and consequent are sets of Boolean items with $X \cap Y = \emptyset$.

A dataset contains a set of N transactions $T = \{t_1, t_2, \ldots, t_N\}$. Each transaction comprises a subset of items, referred to as an itemset, from M items $I = \{i_1, i_2, \ldots, i_M\}$. To extract association rules from datasets, the support-confidence framework was introduced with the Apriori algorithm by Agrawal and Srikant (1994). The support count of an itemset measures the number of transactions that contain an itemset. The support count for an itemset, $\sigma(X)$, is defined as

$$\sigma(X) = |\{t_i | X \subseteq t_i, t_i \in T\}|. \qquad (1)$$

The support, s, determines the strength of a relationship by measuring how often the rule occurs in a dataset according to

$$s(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{N}. \tag{2}$$

The confidence, c, determines how frequently the items in the consequent occur in transactions containing the antecedent, which measures the reliability of the inference, and is defined as

$$c(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)}. \tag{3}$$

These measures have minimum thresholds that are used by a deterministic method to extract rules from the dataset.

2.1. Temporal association rule mining. The term temporal association rule has an ambiguous meaning because it can cover a broad area of temporal data mining (Mitsa, 2010). This paper focuses on discovering association rules where the frequency of their occurrence (i.e., support) changes throughout a temporal dataset, which is a different temporal pattern to that of sequential rules (Agrawal and Srikant, 1995).

Temporal patterns in association rules can be of great benefit to businesses. For example, interesting temporal patterns may coincide with large sports events (e.g., Saleh and Masseglia, 2010) or when unforeseen events occur, such as hurricanes (e.g., Leonard, 2005). As suggested by Yoo and Shekhar (2009), this knowledge can be used to improve the operation of supply chains and the visibility of items during key times. The importance of this problem lies in the ability to discover and adapt to changes with well-informed information. Many domains can benefit from knowledge that has a temporal aspect; within business timely actions are a key to success, whilst for scientific applications it stands to enhance insight and understanding. A review of association rules by Laxman and Sastry (2006) shows that incorporating temporal information can realise greater descriptive and inferential power. It can also provide an additional aspect of interestingness (Freitas, 2002).

Methods based on the support-confidence framework discard rules that fall below the minimum support threshold. Some of these rules can have a relatively high support over a short period of time so these are temporal association rules. The *lifespan* property (Ale and Rossi, 2000) for Boolean association rules is an example that measures support across a subset of a dataset from when the items were first made available to when they stopped being available, or taken off the shelf. This captures an element of the dynamic nature of a dataset because the support measure is not static across the dataset. For example, some supermarket items may be sold only during particular seasonal periods, resulting in annual support values dropping below a minimum threshold, despite having sufficient support values in a seasonal period. The *lifespan* property does not consider datasets where the frequency of rules may increase in temporal periods, whilst still occurring throughout the entire dataset.

Cyclic association rules (Özden *et al.*, 1998) have support measures that change during an item's lifespan and these rules can analyse fluctuations of nonseasonal items. The dataset is partitioned to a desired time granularity and rules are induced from each partition. Cyclic patterns are identified by representing support values of association rules in the partitions as binary sequences and performing pattern matching. These rules are considered to be fully periodic because they repeatedly occur at regular intervals.

Partially periodic rules (Han *et al.*, 1998) do not have the same consistent regularity that is found in fully periodic ones. The cyclic behaviour is found in only some periods of the dataset, and it is not repeated regularly to allow some cycles to be skipped.

Calendar-based schemas (Li *et al.*, 2003) define the temporal intervals for discovering association rules so the requirement of prior knowledge is reduced. A calendar pattern is defined from calendar units such as day, week and month. The Apriori algorithm was modified to discover temporal patterns using calendar-based schemas.

These works illustrate the types of temporal patterns that can potentially be extracted with our proposed method. The importance of temporal patterns in association rules has been discussed with some real-world examples. The temporal feature of the patterns that this paper discovers is one aspect of what makes this a challenging problem.

2.2. Fuzzy association rule mining. In addition to the temporal aspect of this problem, the area of fuzzy association rule mining is relevant. Association rule mining represents the occurrence of items with Boolean attributes. Quantitative association rule mining extends this by discovering rules in quantitative attributes of a dataset. Srikant and Agrawal (1996) introduced mining quantitative association rules by analysing correlations between the intervals of item attributes. The quantitative values are first discretised uniformly and each discretised interval is mapped to an item in the mining process.

Maeda *et al.* (1995) introduced the idea that fuzzy sets can enhance the interpretability and improve robustness to noise for knowledge representation in data mining. Fuzzy association rules were introduced to express quantitative attributes with linguistic labels in a way that is more natural to human reasoning. Chan and Au (1997) proposed F-APACS for mining fuzzy association rules, and this improved the representation with inaccuracies in physical measurements. These are linguistic approaches that provide a linguistic

interpretation of numerical values by experts. Fuzzy association rules also deal with unnatural boundaries of crisp intervals (Kuok *et al.*, 1998). The crisp boundary problem (also referred to as the sharp boundary problem) exists because an attribute may occur near a boundary of a crisp interval but still has the same set membership as if it were in the middle of a crisp interval. Fuzzy sets overcome this because they provide variability in set membership near the boundaries of fuzzy sets. Hong and Lee (2008) reviewed fuzzy association rule mining.

Defining crisp intervals of attributes from this discretisation process, and also partitioning with fuzzy sets, can cause information to be lost (Mata *et al.*, 2002). A loss of information can occur because rules are generated from the intervals and so these rules are totally dependent on how the intervals are constructed. Soft computing techniques can overcome this issue; for example, in the work of Mata *et al.* (2002), a GA evolves attribute intervals to produce rules. Evolutionary algorithms, such as GAs, are suitable for association rule mining because they can search complex spaces and address difficult optimisation problems, which has led to much recent interest in this data mining problem (Herrera, 2008). A review of evolutionary algorithms in association rule mining was conducted by del Jesus *et al.* (2011).

Evolutionary algorithms are often used for discovering intervals specific to rules (Mata *et al.*, 2002). However, when representing attributes with fuzzy sets, the method often used is to define the membership functions first and then mine the rules with an exhaustive search process. Typical methods for defining membership functions for fuzzy association rules are clustering (e.g., Kaya and Alhajj, 2003), expert knowledge (e.g., Chan and Au, 1997) and GAs (e.g., Hong *et al.*, 2008). Simultaneously evolving both membership functions and fuzzy rules in FRBSs is particularly suitable for FRBS controllers (Homaifar and McCormick, 1995) and approximate fuzzy association rules (Matthews *et al.*, 2011b; 2011a). In these works, the purpose of simultaneously evolving both the definition of membership functions and rule induction focuses on improving accuracy (accuracy is discussed in more detail in the next section). Previous approaches for discovering temporal patterns in fuzzy association rules (Matthews *et al.*, 2011b; 2011a) use an approximate FRBS approach because the focus on accuracy allows the discovery of hidden temporal patterns.

There are some works that are close to the problem that this paper addresses, but they are not the same. Martínez-Ballesteros *et al.* (2011) mine a multi-dimensional time series for temporal patterns of association rules. Lee and Lee (2004) used fuzzy sets to describe calendar expressions of temporal patterns. Weng (2011) investigated the temporal property of not discovering rare fuzzy itemsets, which is different to our research, because we focus on how the fuzzy sets are defined instead of only the temporal property. Matthews *et al.* (2011b; 2011a) focused on the composition of fuzzy and temporal association rule mining, which is extended in this paper.

3. Approximate approach

There are two different approaches to using fuzzy sets for representing numeric values in a Mamdani FRBS. The choice of an approximate Mamdani FRBS approach over a linguistic Mamdani FRBS approach is discussed here. It should be noted that the approach in this paper is not a FRBS because it is has no inferencing mechanism but is built on the same principles as FRBSs.

For representing numeric values in a Mamdani FRBS there is the linguistic approach that labels fuzzy sets with linguistic labels to allow an expert to interpret their meaning (e.g., Chan and Au, 1997; Kuok *et al.*, 1998). An approximate FRBS approach does not label the fuzzy sets with linguistic labels (e.g., Homaifar and McCormick 1995; Matthews *et al.*, 2011b; 2011a). Instead, fuzzy numbers are used (Klir *et al.*, 1997). These approaches have a well-known compromise between the interpretability and accuracy of using linguistic or approximate FRBS approaches, and that makes this an important decision. Interpretability can be classified as either complexity-based or semantic-based (Gacto *et al.*, 2011), and this paper states the type of interpretability when it is referred to. Complexity-based interpretability aims to reduce the complexity of a model (e.g., the number of association rules) and semantic-based interpretability maintains the semantics of a model so it is understandable by humans.

It was highlighted by Cordón *et al.* (2001, p. 16) that a Mamdani FRBS using linguistic labels has a rigid partitioning of the membership functions representing the input and output spaces, and this limits the flexibility. It is this lack of flexibility that can prevent some temporal rules from being discovered because there may be rules that occur at the intersection of these membership function boundaries, which were defined before the mining process.

If a linguistic approach were to be used, it would require numerous linguistic labels to achieve the desired level of granularity for discovering all temporal patterns. As a consequence of more linguistic labels, the number of rules increases and so the complexity-based interpretability is reduced (Cordón *et al.*, 2001, p. 16), which counteracts the purpose of a linguistic approach.

Approximate Mamdani FRBSs overcome these problems by allowing some flexibility in rule structure (Cordón *et al.*, 2001). This flexibility allows an increase in the degrees of freedom and expressiveness

that the rules can take because each rule has its own distinct membership functions. Flexibility is crucial for discovering temporal patterns because it allows rules to be discovered that would have previously fallen on the intersections of membership function boundaries. This is an important factor in choosing an approximate FRBS approach. Such an approach is good for expressing temporal patterns because it provides a mechanism to model the fine detail of temporal patterns that are otherwise lost with a linguistic approach. A flexible approach to the definition of fuzzy sets also allows complex problems to be modelled with as many rules as required. This means that many temporal patterns with many different fuzzy sets can be discovered.

However, the flexibility of an approximate FRBS approach creates a lack of semantic-based interpretability, because the membership functions are not assigned linguistic labels. Despite this, it is a step towards discovering these patterns that would otherwise be lost. Assigning linguistic labels to fuzzy sets can cause a loss of rules, because the fuzzy sets are inflexible and they have low membership functions where their boundaries intersect with other fuzzy sets. So, the accuracy of an approximate FRBS approach is required to discover these temporal patterns.

Although the semantic-based interpretability is lost because there are no linguistic labels, efforts have been made to enhance the complexity-based interpretability of the shape of membership functions. An objective of the evolutionary algorithm (see Section 4.3.3) is to restrict the membership function's shape to an ideal width so that its semantic-based interpretability is increased. This is considered an improvement in semantic-based interpretability because the shape of the membership function is a factor in interpreting its meaning. For example, a very thin membership could cover such a small section of its domain that it becomes difficult to ascertain its meaning. This objective provides a mechanism for preventing the membership functions from evolving to cover the entire input domain, whilst allowing enough accuracy to represent a temporal pattern.

As seen from the literature in Section 2.2, there are approximate FRBS approaches for mining quantitative association rules using crisp intervals that are not defined before the mining process (e.g. Mata *et al.*, 2002). Representing the quantitative attributes with fuzzy sets extends other approaches by better handling uncertainty and overcoming the crisp boundary problem.

4. Multi-objective evolutionary search and optimisation

A MOEA was applied to this problem because of several key features of this class of algorithms that are discussed here.

Many problems have two or more objectives, often competing, where a trade-off between these objectives is required to solve a problem. A MOEA is a method for solving this that uses populations of candidates solutions to explore many different possible trade-offs. A common feature of many MOEAs is the concept of Pareto dominance. An objective vector of a solution is said to dominate the objective vector of another solution if no component is smaller and at least one is greater than the components of the other solution's objective vector (Deb, 2005). In other words, a solution is said to be Pareto optimal when no change in the solution will improve one objective without degrading another objective. The set of Pareto optimal solutions is known as the Pareto set and the corresponding set of objectives for those solutions is known as the Pareto front.

The decision variable space for this problem contains variables for the items, the membership functions and the temporal periods. It is a challenging problem to simultaneously search this space because the membership functions for rules can be different for different temporal periods of the dataset. There are multiple temporal patterns that can be present in the dataset and a MOEA provides a method for discovering some of these rules. It is likely there are numerous trade-off solutions so the Pareto set will have a cardinality greater than one. The variety in the Pareto set is important because it is not the final solution itself; instead, the Pareto set is used by an expert in decision making (Hopgood, 2012). This applies to association rule mining because not all rules discovered may be interesting to an expert, despite the strength in interestingness measures. Matthews *et al.* (2011b) evolve one rule in each run of a GA and this work (also Matthews *et al.*, 2011b) extends that with a MOEA to yield several rules from one run of a GA. Pareto dominance creates numerous trade-off solutions for the multi-objective problem and this feature was used to yield multiple temporal patterns. IRL also produced multiple rules and this is discussed in detail in Section 4.2.

The support and confidence measures are competing objectives (Bayardo and Agrawal, 1999) and MOEAs are often used to mine association rules because of their ability to handle competing objectives. Previous MOEAs for association rule mining have focused on Boolean data (e.g., Ghosh and Nath, 2004) and quantitative datasets (e.g., Miller and Yang, 1997), but not the composition of temporal and quantitative, which is a significant step in problem dimensionality. From the plethora of MOEAs, the Nondominated Sorting Genetic Algorithm II (NSGA-II) (Deb *et al.*, 2002) was selected for its popularity (Alcalá *et al.*, 2007a) and ability to maintain a diverse set of solutions suitable for extracting multiple patterns. Previous works have used the NSGA-II for subgroup discovery (Carmona *et al.*, 2010), a closely related area, motif sequence discovery (Kaya, 2009), a

different form of temporal mining, and also predictive tasks (Ben Aicha *et al.*, 2013).

4.1. Representation. A chromosome of mixed types represents an entire association rule. A Michigan approach and a mixed coding scheme were used to represent the temporal period and fuzzy association rules. A chromosome, C, was defined as

$$C = (e_l, e_u, i_1, a_1, b_1, c_1, A_1, \ldots, i_k, a_k, b_k, c_k, A_k),$$
(4)

where the lower temporal endpoint is e_l, the upper temporal endpoint is e_u, items are integers denoted with i (e.g., pizza), the basic parameters of the asymmetrical triangular membership functions are real numbers indicated with a, b and c for association rules with k items and a binary value in A determines if the clause belongs to the antecedent or consequent part of the rule. Unconstrained learning (Cordón and Herrera, 1997) is applied to the membership functions where no restrictions are placed on intervals except that $a \leq b \leq c$. All parts are randomly initialised according to the following: each item i is distinct, the triangular membership function parameters are $a \leq b \leq c$, and at least one antecedent and one consequent are present in a rule. Additionally, the minimum temporal support (Matthews *et al.*, 2011b; 2011a) restricts the temporal distance between the endpoints based on

$$e_u - e_l \geq \text{minTempSup}.$$
(5)

The minimum temporal support defines the smallest granularity of a temporal period that can be evolved. This is similar to a unit representation in a calendar-based schema because the length of the temporal period is defined. But, it is also different because the temporal period can start and end at any transaction, which allows flexibility in where the temporal period occurs in the dataset. Without minimum temporal support, solutions evolve to cover very few transactions because this yields a stronger support value.

4.2. Iterative rule learning. IRL is a process where one rule is chosen from one run of a GA. The GA is run many times to produce a set of rules. This is an approach first used in the work of Venturini (1993), where a GA continues to extract classification rules when there are still examples labelled "uncovered". In this paper, IRL was performed by maintaining a set of final rules evolved from each run of the GA. This final rule set contains all discovered rules and is considered the final result of this data mining method. A MOEA produces many solutions as the final output of one run and so only the rules from the Pareto front are used in IRL because these show the best trade-off solutions. Figure 2 illustrates the process of IRL with a MOEA.
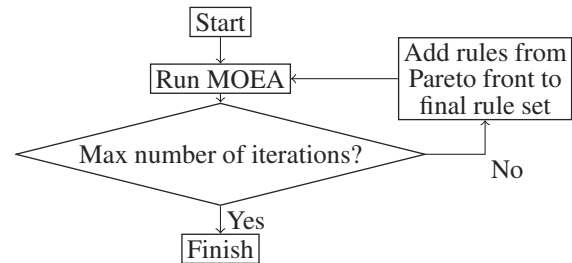

Fig. 2. Process of iterative rule learning.

Chromosomes are penalised in the fitness function if the candidate rule matches a rule from the final rule set, similarly as in the work of Mata *et al.* (2002). A match was determined based on the following criteria:

- The same clauses of the rule. The parts of a clause that are compared are the items, the gene that identifies the consequent or antecedent part of the rule, and the fuzzy sets. The following similarity measure, based on intersection, union and cardinality (Dubois and Prade, 1980, p. 24), compares two fuzzy sets, E and F,

$$S(E, F) = \frac{\|(E \cap F)\|}{\|(E \cup F)\|},$$
(6)

where the notation '$\|\cdot\|$' represents scalar cardinality. If the similarity of the two fuzzy sets is greater than a threshold, then the clauses are considered to be a match.

- The temporal endpoints have an overlap above a threshold.

Candidate chromosomes that have previously been discovered, because they are present in the final rule set, are penalised by setting their fitness to 0. This penalisation method helps to guide search away from previously discovered rules so that the final rule set is diverse. IRL was terminated after a set number of iterations.

4.3. Objectives. The fitness objectives are designed to search the multi-dimensional objective space. Temporal fuzzy support and confidence measure how good a rule is, but they are not used as thresholds in this approach. The following objectives are minimised to zero.

4.3.1. Temporal fuzzy support. The temporal fuzzy support objective guides evolution towards discovering rules that occur more frequently in temporal periods of the dataset. Modified from the work of Ale and Rossi (2000)

to include a fuzzy representation, this was redefined to a minimisation function as

$$\text{TemporalFuzzySupport}(C) \tag{7}$$
$$= 1 - \left(\frac{\sum_{j=e_l}^{e_u} \text{FuzzySupport}(C_X^{(j)} \cap C_Y^{(j)})}{e_u - e_l} \right),$$

with the same definitions as before: X and Y are itemsets, j is a dataset transaction, and e_l and e_u are endpoints. The FuzzySupport (Hong *et al.*, 2001) uses minimum for the intersection operation.

The minimum temporal support prevents solutions evolving towards the smallest time interval that produces the highest temporal support, e.g., covering 1 transaction would give a temporal support of 100 (or 0 in the case of this minimisation objective).

4.3.2. Confidence. The confidence identifies how common the antecedent is among other transactions. This was redefined to a minimisation function as

$$\text{Confidence}(C) \tag{8}$$
$$= 1 - \left(\frac{\sum_{j=e_l}^{e_u} \text{FuzzySupport}(C_X^{(j)} \cap C_Y^{(j)})}{\sum_{j=e_l}^{e_u} \text{FuzzySupport}(C_X^{(j)})} \right).$$

4.3.3. Membership function widths. The aim of this objective was to prevent the membership function parameters from evolving to cover the entire attribute domain. Covering the entire domain would yield a high fuzzy support value for a rule because the rule would then be present in many transactions. The Root Mean Square Error (RMSE) measures the difference between the membership function widths and an ideal width by

$$\text{RMSE} = \sqrt{\frac{1}{k} \sum_{j=1}^{k} (w - (c_j - a_j))^2}, \tag{9}$$

where the ideal width is w. The purpose of the ideal width was to prevent solutions evolving towards a width of 0 that would be detrimental to the semantic-based interpretability of the membership function.

4.4. Genetic operators. The crossover operator consists of four separate methods for operating on different sections of the chromosome that have their own constraints.

The section of a chromosome containing the lower and upper endpoints are crossed over with uniform crossover and the feasibility of offspring was ensured by

satisfying the minimum temporal support constraint on endpoints e_l and e_u, see Eqn. (5).

For the itemsets found in the next section of a chromosome, uniform crossover was adapted to ensure that the offspring do not contain duplicate items. The method of crossover of itemsets by Matthews *et al.* (2010) was extended to include membership functions. The advantage of this method is that the ordering of items remains unless a duplicate is present in the itemset. A summary of each stage of the crossover is briefly described here.

Stage 1.
Merge the chromosomes from two selected parents into an intermediate array so that no two items from the same parent are adjacent.

Stage 2.
Check each item in the array for duplicate values against the remaining items. If a duplicate is found, the duplicate item is swapped with the next item. The result is that all duplicate items are adjacent and the items can now be selected from the intermediate array to form offspring.

Stage 3.
Select items from the intermediate array by iterating over every even index value. A random integer (0 or 1) is added to the index and the indexed item is added to the offspring. If a 0 is generated, it is checked for duplicates with the preceding item, and if a duplicate is found, it adds 1 to the index, otherwise it adds 0.

Membership functions only have crossover performed when the items are the same, for example, crossover of membership functions is performed on items pizza and pizza, but not for items pizza and beer. This prevents crossing over membership function parameters from different items that would lose the semantics of the membership functions and it would be more exploratory (i.e., mutation) than exploitative. For items that do not match, the membership functions remain unchanged by copying them across to the offspring. For matching items, the parent centric BLX-α (PBX-α) (Lozano *et al.*, 2004) crossover operator was applied to the continuous attributes of the membership functions parameters. The PBX-α crossover operator has a higher probability of producing offspring within the neighbourhood of its parents, rather than the entire search space. The α parameter controls the diversity that this operator introduces to a population.

The gene that determines whether the clause is part of the antecedent or consequent was randomly swapped. This was only swapped for items that match; the same method that was used for the membership functions was applied to this gene as well.

Mutation operates in an identical manner to initialisation. A gene was chosen at random and its allele is randomly created whilst ensuring it is feasible.

5. Methodology

The methodology is described for analysing the ability of the approximate FRBS approach for discovering temporal fuzzy association rules for single and multiple patterns. To demonstrate the efficacy of our proposed approach, several experiments were conducted on datasets with different features. The datasets are introduced first, followed by the methodology, and then the parameters' settings for the MOEA are given.

5.1. Dataset generation. The IBM Quest Synthetic Data Generator[1] (Agrawal and Srikant, 1994) was extended to include quantitative attributes based on a method similar to that of Hong *et al.* (2008) with a simulated dataset. This generates market basket datasets, and it can be considered a benchmark because it has been used in many works (e.g., Agrawal and Srikant, 1994; Özden *et al.*, 1998; Li *et al.*, 2003; Matthews *et al.*, 2011b; 2011a). The generated datasets are some of very few datasets that have time-stamped transactions, quantities of items, and a variable number of items in each transaction (variable size shopping baskets are unlike time-series data). The quantities of items sold in a transaction were randomly generated from a uniform distribution in the range 1–20 and assigned to each item in a transaction. The dataset had these features: 1000 transactions, 50 items and an average transaction size of 10. This dataset was used as the basis for all experiments conducted.

Each transaction was identified with a transaction ID and this was assumed to be recorded in the order that transactions occur so the dataset can be considered temporal. This assumption has also been made in other works focusing on temporal data mining (e.g., Yoo and Shekhar, 2009).

5.2. Augmentation. Temporal association rules were augmented into datasets to provide a controlled experiment for assessing our proposed approach. The augmented patterns act as targets, and the aim of our approach was to discover these target patterns. The augmented patterns are single points on the universe of discourse. These points are intended to represent points that can occur at the intersection of membership function boundaries for a descriptive approach. Discovering the augmented patterns with fuzzy numbers (approximate

approach) can demonstrate how the loss of rules can be avoided. It should be noted that a comparative analysis with existing approaches, which use a descriptive FRBS approach, is not performed, because the aim of our work is to assess the ability of an approximate FRBS approach with a MOEA in a complex search space. The augmentation method is now described.

The Apriori algorithm was used to identify Boolean association rules. The minimum support was set to 2.5% and minimum confidence set to 50%. Boolean association rules were chosen from the results of the Apriori algorithm, and each rule was augmented in a new copy of the dataset. The criteria for choosing rules are described in the next paragraph. After selecting a rule, augmentation occurs by first identifying the transactions that contain the chosen rule. A temporal period of the dataset was chosen that has an arbitrary number of these transactions containing the rule. This temporal period was within the smallest temporal granularity defined by the minimum temporal support. These transactions are then modified to augment a temporal pattern in that temporal period. The modification occurs by changing the quantitative attribute of the items in a transaction, but only for those items that match the rule. The quantitative attributes were modified to be the same in the temporal period. The result is a dataset that has an increase in the quantitative association rule in a temporal period. It is this quantitative association rule that will be represented by a fuzzy association rule.

An example dataset that has been augmented is demonstrated in Table 1. The augmented temporal pattern occurs in transactions 5 and 6 of the example dataset. The quantitative values were 5 and 15. These values were chosen as an example of the location of the intersection of membership function boundaries if the descriptive approach of Fig. 1 were used. A rule at the intersection of membership function boundaries has a low fuzzy support, so the intention of augmenting a temporal pattern is to replicate this situation in order to analyse whether the MOEA can discover the rule.

The augmentation method uses the Apriori algorithm to find relatively low (3.1%), medium (4.9%) and high support (5.7%) Boolean association rules. The different strength Boolean association rules were used to produce datasets with various levels of difficulty. The reason for this is that the itemset space is one of the more important search spaces that the MOEA covers because a rule must match one transaction before fuzzy support can be calculated. Datasets with various levels of difficulty allow the efficacy to be assessed. Rules to augment were chosen manually.

This method augments a temporal pattern from existing transactions. It differs from other augmentation methods that insert an entirely new transaction, such as those by Matthews *et al.* (2011b; 2011a). Augmenting an

---

[1]Original source code no longer available. See alternative, `http://www.cs.nmsu.edu/~cgiannel/assoc_gen.html` (last accessed 5th December 2010).

Table 1. Example of how a temporal pattern was augmented into a dataset containing market basket transactions. Transactions with a strikethrough line show the transaction before it was modified. The items in bold are the augmented pattern.

TID	Items
1	{Bread(12), Pizza(2), Beer(1)}
2	{Eggs(6), Milk(1)}
3	{Eggs(2), Pizza(4)}
4	{Bread(4), Cheese(18)}
~~5~~	~~{Cheese(2), Pizza(1), Beer(12)}~~
5	{Cheese(2), **Pizza(5)**, **Beer(15)**}
~~6~~	~~{Pizza(2), Beer(4), Milk(18)}~~
6	{**Pizza(5)**, **Beer(15)**, Milk(18)}
7	{Eggs(18), Cheese(3)}
8	{Bread(3), Eggs(12), Cheese(1)}
9	{Pizza(1), Cheese(9)}
10	{Bread(1), Pizza(11)}

existing transaction means that the combinatorial search space for items is not altered, and so the search is not made easier by inserting new transactions. Figure 3 shows the sum of fuzzy support of the augmented pattern before and after augmentation of a dataset. The augmented pattern is the large peak and this is the target pattern for the MOEA to discover. Table 2 shows the augmented datasets and the naming convention Dqp is adopted where q is the quantity of temporal patterns (1 in the case of Table 2) and p describes the augmented pattern (L/M/H for low/medium/high support rules, respectively).
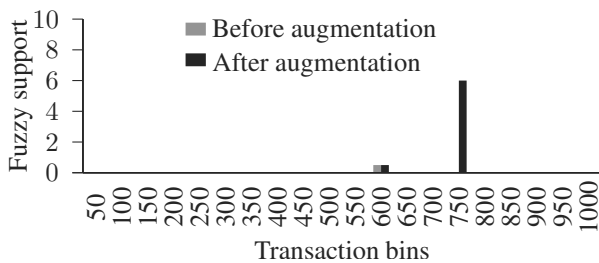

Fig. 3. Bins (size 50) of fuzzy support for an augmented pattern in D1M.

Table 2. Description of datasets augmented with 1 temporal pattern. Support and confidence measure the rules before they were augmented.

Data-set	Sup. (%)	Conf. (%)	Endpoint e_l	Endpoint e_u	Quantitative rule
D1H	5.7	53.3	700	746	17(5) $\Rightarrow$ 21(15)
D1M	4.9	51	649	685	22(5) $\Rightarrow$ 8(15)
D1L	3.1	50.8	566	591	34(5) $\Rightarrow$ 45(15)

5.3. Multiple rules.

More datasets were created using the same augmentation method but for two patterns in different temporal periods of the dataset. The aim of augmenting two temporal patterns was to analyse the efficacy of our proposed method at discovering multiple rules. Our proposed method can produce multiple rules from both the Pareto front of the MOEA and also with numerous iterations of the MOEA using IRL.

The temporal patterns are augmented so that one pattern appears earlier in the dataset than the next pattern, so they do not overlap. The number of modified transactions was identical in both augmented patterns and the length of the temporal period was kept the same as much as possible in the dataset. This ensures that the temporal fuzzy support yields values that are approximately the same for each augmented pattern. Consequently, both patterns were global minima that were used to analyse the efficacy in discovering multiple rules. The datasets are described in Table 3 and had various levels of difficulty for the purpose of analysing efficacy. An illustration of the temporal patterns is provided in Fig. 4.

Table 3. Description of datasets augmented with 2 temporal patterns. Support and confidence measure the rules before they were augmented.

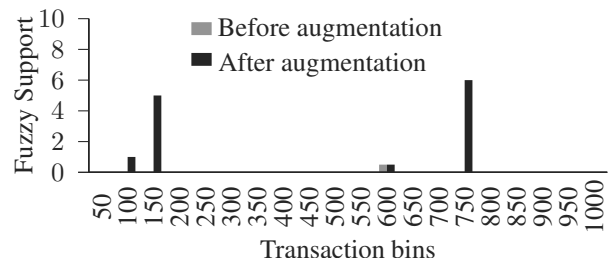| Data-set	Sup. (%)	Conf. (%)	Endpoint e_l	Endpoint e_u	Quantitative rule
D2H	5.7	53.3	112	155	17(5) $\Rightarrow$ 21(15)
			700	746	17(5) $\Rightarrow$ 21(15)
D2M	4.9	51	128	166	22(5) $\Rightarrow$ 8(15)
			752	791	22(5) $\Rightarrow$ 8(15)
D2L	3.1	50.8	566	591	34(5) $\Rightarrow$ 45(15)
			791	820	34(5) $\Rightarrow$ 45(15)

Fig. 4. Bins (size 50) of fuzzy support for an augmented pattern in D2M.

5.4. Parameter settings.

Several parameters were fixed, because the intention was to demonstrate the MOEA for this data mining problem. The NSGA-II was configured with the following properties: population size of 1000, 500 generations, 40% crossover, 20% mutation,

90% for α in PBX-α, a rule length of 2, and the ideal width of membership functions was set to 2. IRL was set to perform 20 runs of the NSGA-II, the fuzzy similarity threshold was set to 0.5, and the overlap of the temporal endpoints was set to 75%. The IRL algorithm was repeated 30 times on each dataset.

The minimum temporal support was set differently for each dataset. When augmenting a temporal pattern, the size of the temporal periods varied with different rules because they had different support values, and so transactions containing the rule appeared at different frequencies in the dataset. The minimum temporal support was set according to the size of the temporal period. For example, in dataset D1M (of Table 2), the temporal pattern starts at transaction 649 and ends at 685, a difference of 36, so the minimum temporal support was set to 40.

6. Results

The method for analysing the results is presented first and the results then follow.

6.1. Method of analysis. The results of each run were analysed to identify if the augmented pattern was discovered. The method for penalising rules in IRL (see Section 4.2) was also used to compare rules in the evaluation. An augmented pattern is considered to be discovered if a fuzzy association rule from the final rule set has the following criteria:

- The same itemsets and similar fuzzy sets. The same method as used for IRL; however, whether the item is in the antecedent or consequent is not included, because only fuzzy support is used to identify the rule, and not confidence.

- Fuzzy support is above a threshold value of 90% for each transaction.

- The temporal endpoints have an overlap above 75%. This criterion is identical to that used in IRL, see Section 4.2.

6.2. Discovering augmented patterns. The results for running the MOEA on a dataset with one augmented pattern are summarised in Table 4. For each dataset, the MOEA successfully discovered the augmented pattern in at least one run of IRL. The results show that our proposed approach was capable of finding temporal patterns in the datasets used because the target patterns were discovered.

The augmented temporal pattern was discovered in all runs of IRL on a dataset containing a strong temporal pattern (D1H). The augmented patterns were discovered less in D1M (27 times) and the target pattern was discovered only twice in D1L. The number of times the

Table 4. Results from 30 runs of IRL on each dataset that has one temporal pattern augmented.

Dataset	Minimum temporal support	Discovered target
D1H	50	30
D1M	40	27
D1L	30	2

augmented patterns were discovered shows the efficacy of our proposed approach with different levels of difficulty. The results indicate that our proposed method was better at discovering patterns in datasets where the quantitative itemset has a high fuzzy support in the dataset. The confidence values of the target patterns discovered by our proposed approach were all high values compared to the variety of confidence values of other rules.

Fitness values for all objectives of a final result for one run with dataset D1M are shown in Fig. 5. The augmented pattern is indicated by a different point style, '+', and this shows the target pattern in the objective space standing out from other rules. There was a trade-off between objectives in the target pattern. Drawing attention to all of the rules plotted, a clear trade-off can be seen between temporal fuzzy support and the width of membership functions. There was less of a trade-off between the confidence and other objectives, but this is a key factor used in determining the interestingness of a rule. Viewing such trade-offs is important when analysing the results of the mining process, particularly for knowledge discovery in real-world applications.
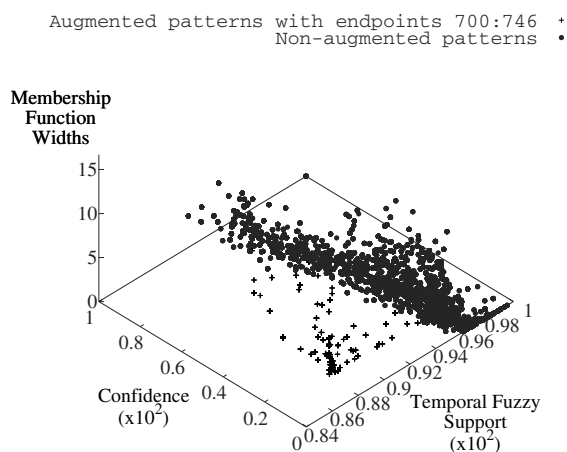

Fig. 5. All rules from one run of IRL on D1H. Each point represents a rule in the objective space.

The membership functions of one rule evolved from dataset D1H are shown in Fig. 6. These membership functions have evolved to the ideal membership function width.
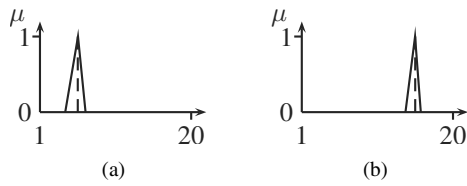
Fig. 6. Membership functions of an augmented rule that was discovered in dataset D1H. The dashed line represents the augmented pattern: item 17 with quantity 5 and membership function parameters $a = 3.34$, $b = 5.03$, and $c = 6.03$ (a), item 21 with quantity 15 and membership function parameters $a = 13.71$, $b = 14.98$, and c 15.74 (b).

The results for datasets with two augmented temporal patterns are summarised in Table 5. The analysis here focuses on the efficacy of IRL to discover multiple patterns, which uses iterations to discover more patterns. The results show that for each dataset, IRL successfully discovered at least one augmented pattern (either 1st or 2nd) in at least one run of IRL. This shows that IRL was capable of discovering patterns that share global minima because augmented patterns were evolved from different temporal periods of the datasets tested.

Table 5. Results from 30 runs of IRL on each dataset that has two temporal patterns augmented. The 1st and 2nd discovered targets refer to the first augmented pattern appearing earlier in a dataset than the second augmented pattern.

Dataset	Minimum temporal support	Discovered targets		
		1st	2nd	Both
D2H	50	30	30	30
D2M	40	17	17	2
D2L	30	2	4	0

The number of IRL runs that discovered both augmented temporal patterns was also analysed. The analysis continues to focus on the the efficacy of IRL to discover multiple patterns. Dataset D2H had the best efficacy for discovering multiple patterns with IRL because it discovered both patterns in every run of IRL whilst D2M discovered both patterns less (2 times) and D2L did not discover them at all. The results show the efficacy of IRL to be good on the datasets tested and with mixed performance on datasets with various levels of difficulty.

The analysis now focuses on the efficacy of the MOEA to discover multiple patterns. Table 6 shows how many times a single run of a MOEA discovered both augmented temporal patterns in all runs of the MOEA in all iterations of IRL. Both patterns were evolved in dataset D2H in 107 of 600 runs of the MOEA, only twice with dataset D2M, and both patterns were not evolved with

any runs in dataset D2L. This demonstrates the ability of the MOEA, the NSGA-II, in maintaining a diverse set of solutions that allows two temporal patterns to simultaneously evolve in one run. All the rules evolved for dataset D2H are plotted in Fig. 7 and the two augmented temporal patterns can be seen. This shows the same trade-off in solutions that was illustrated in Fig. 5.

Table 6. Results from all runs of the MOEA in each run of IRL (total of 600) on each dataset that has two temporal patterns augmented.

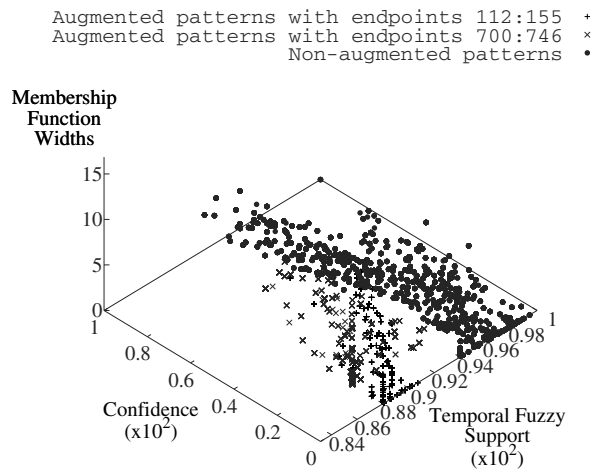| Dataset	Minimum temporal support	Discovered both targets
D2H	50	107
D2M	40	2
D2L	30	0

Fig. 7. All rules from IRL for D2H. Each point represents a rule in the objective space.

7. Conclusions

This paper has proposed a MOEA for discovering fuzzy association rules that have a temporal pattern. The approach uses an approximate FRBS method to overcome the issues arising from defining the membership functions before the mining process. The experimental methodology has augmented temporal patterns with quantitative values, which would be near to the intersection of membership function boundaries in a descriptive FRBS approach. This has allowed the ability and efficacy of our proposed method to be analysed.

The main finding is that our proposed method is capable of discovering temporal patterns of fuzzy association rules of the datasets tested. This contributes towards discovering patterns that could otherwise be

lost from defining the membership functions before the mining process. Through using datasets with various levels of difficulty, it was shown how the efficacy of discovering temporal patterns is affected by the support of its Boolean itemset. A temporal pattern that has a Boolean itemset, where the itemset occurs frequently, is more likely to be discovered than an itemset that is less frequent. Future work will address mining temporal fuzzy association rules that have lower fuzzy support across the whole dataset.

The combination of a MOEA and IRL has provided a method for mining multiple rules. This is particularly useful for the knowledge discovery process because it provides a variety of rules for an expert to determine which ones are meaningful and useful within the context of a particular application. IRL has greatly helped in discovering the temporal patterns that are considered to be more difficult as it provides a completely new run of the algorithm that is guided away from previously discovered solutions.

The method is proposed as a process to further support the data mining process and not replace existing methods. The reason for this is that it can find patterns that other methods may not as a result of defining the membership functions before the mining process.

Having demonstrated the ability to find target patterns that have been defined by a methodology of augmentation, the natural progression of the methodology is to perform a comparative analysis with existing approaches. There is potential for future work to explore different types of temporal pattern and different fuzzy set representations that provide flexibility. Furthermore, flexibility may be suitable for non-temporal problems. The robustness of identifying quantitative attributes and evolving low support itemsets will be addressed. Real-world datasets will be used and an analysis of scalability will be conducted. Theoretical analysis will be conducted, which can provide further explanation of our empirical analysis.

Acknowledgment

References

Agrawal, R., Imieliński, T. and Swami, A. (1993). Mining association rules between sets of items in large databases, *Proceedings of the ACM SIGMOD International Conference on Management of Data, Washington, DC, USA*, pp. 207–216.

Agrawal, R. and Srikant, R. (1994). Fast algorithms for mining association rules, *Proceedings of the 20th International Conference on Very Large Data Bases, Santiago, Chile*, pp. 487–499.

Agrawal, R. and Srikant, R. (1995). Mining sequential patterns, *Proceedings of the 11th International Conference on Data Engineering, Taipei, Taiwan*, pp. 3–14.

Alcalá, R., Alcalá-Fdez, J., Gacto, M. and Herrera, F. (2007a). A multi-objective evolutionary algorithm for rule selection and tuning on fuzzy rule-based systems, *Proceedings of the IEEE International Fuzzy Systems Conference (FUZZ-IEEE 2007), London, UK*, pp. 1–6.

Alcalá, R., Alcalá-Fdez, J. and Herrera, F. (2007b). A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection, *IEEE Transactions on Fuzzy Systems* **15**(4): 616–635.

Ale, J.M. and Rossi, G. H. (2000). An approach to discovering temporal association rules, *Proceedings of the 2000 ACM Symposium on Applied Computing (SAC'00), Como, Italy*, pp. 294–300.

Bayardo, Jr., R.J. and Agrawal, R. (1999). Mining the most interesting rules, *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA*, pp. 145–154.

Ben Aicha, F., Bouani, F. and Ksouri, M. (2013). A multivariable multiobjective predictive controller, *International Journal of Applied Mathematics and Computer Science* **23**(1): 35–45, DOI: 10.2478/amcs-2013-0004.

Carmona, C., González, P., del Jesus, M. and Herrera, F. (2010). NMEEF-SD: Non-dominated multiobjective evolutionary algorithm for extracting fuzzy rules in subgroup discovery, *IEEE Transactions on Fuzzy Systems* **18**(5): 958–970.

Casillas, J., Cordón, O., del Jesus, M. and Herrera, F. (2005). Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction, *IEEE Transactions on Fuzzy Systems* **13**(1): 13–29.

Chan, K.C.C. and Au, W.-H. (1997). Mining fuzzy association rules, *Proceedings of the 6th International Conference on Information and Knowledge Management, Las Vegas, NV, USA*, pp. 209–215.

Cordón, O. and Herrera, F. (1997). Identification of linguistic fuzzy models by means of genetic algorithms, *in* D. Driankov and H. Hellendoorn (Eds.), *Fuzzy Model Identification. Selected Approaches*, Springer-Verlag, Heidelberg, pp. 215–250.

Cordón, O., Herrera, F., Hoffmann, F. and Magdalena, L. (2001). *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*, Advances in Fuzzy Systems—Applications and Theory, World Scientific, Singapore.

Deb, K. (2005). Multi-objective optimization, *in* E.K. Burke and G. Kendall (Eds.), *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, Springer, Berlin, pp. 403–449.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation* **6**(2): 182–197.

del Jesus, M.J., Gámez, J.A., González, P. and Puerta, J.M. (2011). On the discovery of association rules by means of evolutionary algorithms, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **1**(5): 397–415.

Dubois, D. and Prade, H.M. (1980). *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, London.

Freitas, A.A. (2002). *Data Mining and Knowledge Discovery with Evolutionary Algorithms*, Springer-Verlag, Berlin/Heidelberg.

Gacto, M., Alcalá, R. and Herrera, F. (2011). Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures, *Information Sciences* **181**(20): 4340–4360.

Ghosh, A. and Nath, B. (2004). Multi-objective rule mining using genetic algorithms, *Information Sciences* **163**(1–3): 123–133.

Han, J., Gong, W. and Yin, Y. (1998). Mining segment-wise periodic patterns in time-related databases, *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, New York, NY, USA*, pp. 214–218.

Herrera, F. (2008). Genetic fuzzy systems: Taxonomy, current research trends and prospects, *Evolutionary Intelligence* **1**(1): 27–46.

Homaifar, A. and 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., Chen, C.-H., Lee, Y.-C. and Wu, Y.-L. (2008). Genetic-fuzzy data mining with divide-and-conquer strategy, *IEEE Transactions on Evolutionary Computation* **12**(2): 252–265.

Hong, T.-P., Kuo, C.-S. and Chi, S.-C. (2001). Trade-off between computation time and number of rules for fuzzy mining from quantitative data, *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems* **9**(5): 587–604.

Hong, T.-P. and Lee, Y.-C. (2008). An overview of mining fuzzy association rules, *in* H. Bustince, F. Herrera and J. Montero (Eds.), *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, Studies in Fuzziness and Soft Computing, Vol. 220, Springer, Berlin/Heidelberg, pp. 397–410.

Hopgood, A. (2012). *Intelligent Systems for Engineers and Scientists*, 3rd edn, CRC Press, Boca Raton, FL.

Kaya, M. (2009). MOGAMOD: Multi-objective genetic algorithm for motif discovery, *Expert Systems with Applications* **36**(2, Part 1): 1039–1047.

Kaya, M. and Alhajj, R. (2003). Facilitating fuzzy association rules mining by using multi-objective genetic algorithms for automated clustering, *Proceedings of the 3rd IEEE International Conference on Data Mining, Melbourne, FL, USA*, pp. 561–564.

Klir, G.J., Clair, U.H.S. and Yuan, B. (1997). *Fuzzy Set Theory: Foundations and Applications*, Prentice Hall, Upper Saddle River, NJ.

Kuok, C.M., Fu, A. and Wong, M.H. (1998). Mining fuzzy association rules in databases, *SIGMOD Record* **27**(1): 41–46.

Laxman, S. and Sastry, P. S. (2006). A survey of temporal data mining, *Sādhanā* **31**(2): 173–198.

Lee, W.-J. and Lee, S.-J. (2004). Discovery of fuzzy temporal association rules, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **34**(6): 111–118.

Leonard, D. (2005). After Katrina: Crisis management, the only lifeline was the Wal-Mart, *FORTUNE Magazine* (October 3).

Li, Y., Ning, P., Wang, X. S. and Jajodia, S. (2003). Discovering calendar-based temporal association rules, *Data & Knowledge Engineering* **44**(2): 193–218.

Lozano, M., Herrera, F., Krasnogor, N. and Molina, D. (2004). Real-coded memetic algorithms with crossover hill-climbing, *Evolutionary Computation* **12**(3): 273–302.

Maeda, A., Ashida, H., Taniguchi, Y. and Takahashi, Y. (1995). Data mining system using fuzzy rule induction, *Proceedings of the 1995 IEEE International Conference on Fuzzy Systems, Yokohama, Japan*, Vol. 5, pp. 45–46.

Martínez-Ballesteros, M., Martínez-Álvarez, F., Troncoso, A. and Riquelme, J. (2011). An evolutionary algorithm to discover quantitative association rules in multidimensional time series, *Soft Computing—A Fusion of Foundations, Methodologies and Applications* **15**(10): 1–20.

Mata, J., Alvarez, J.L. and Riquelme, J.C. (2002). An evolutionary algorithm to discover numeric association rules, *Proceedings of the 2002 ACM Symposium on Applied Computing, Madrid, Spain*, pp. 590–594.

Matthews, S.G., Gongora, M.A. and Hopgood, A.A. (2010). Evolving temporal association rules with genetic algorithms, *in* M. Bramer, M. Petridis and A. Hopgood (Eds.), *Research and Development in Intelligent Systems XXVII*, Springer, London, pp. 107–120.

Matthews, S.G., Gongora, M.A. and Hopgood, A.A. (2011a). Evolving temporal fuzzy association rules from quantitative data with a multi-objective evolutionary algorithm, *in* E. Corchado, M. Kurzynski and M. Wozniak (Eds.), *Hybrid Artificial Intelligent Systems (Proceedings of HAIS 2011)*, Lecture Notes in Computer Science, Vol. 6678, Springer, Berlin/Heidelberg, pp. 198–205.

Matthews, S.G., Gongora, M.A. and Hopgood, A.A. (2011b). Evolving temporal fuzzy itemsets from quantitative data with a multi-objective evolutionary algorithm, *IEEE 5th International Workshop on Genetic and Evolutionary Fuzzy Systems (GEFS 2011), Paris, France*, pp. 9–16.

Miller, R.J. and Yang, Y. (1997). Association rules over interval data, *ACM SIGMOD Record* **26**(2): 452–461.

Mitsa, T. (2010). *Temporal Data Mining*, CRC Press Online, Boca Raton, FL.

Özden, B., Ramaswamy, S. and Silberschatz, A. (1998). Cyclic association rules, *Proceedings of the 1914 International Conference on Data Engineering, Orlando, FL, USA*, pp. 412–421.

Piatetsky-Shapiro, G. (1990). Knowledge discovery in real databases: A report on the IJCAI-89 workshop, *AI Magazine* **11**(4): 68–70.

Saleh, B. and Masseglia, F. (2010). Discovering frequent behaviors: Time is an essential element of the context, *Knowledge and Information Systems* **28**(2): 1–21.

Srikant, R. and Agrawal, R. (1996). Mining quantitative association rules in large relational tables, *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, Quebec, Canada*, pp. 1–12.

Venturini, G. (1993). SIA: A supervised inductive algorithm with genetic search for learning attributes based concepts, *Proceedings of the European Conference on Machine Learning (ECML–93), Vienna, Austria*, pp. 280–296.

Weng, C.-H. (2011). Mining fuzzy specific rare itemsets for education data, *Knowledge-Based Systems* **24**(5): 697–708.

Yoo, J.S. and Shekhar, S. (2009). Similarity-profiled temporal association mining, *IEEE Transactions on Knowledge and Data Engineering* **21**(8): 1147–1161.

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

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning, Parts I, II, III, *Information Sciences* **8–9**(3,4,1): 199–249, 301–357, 43–80.
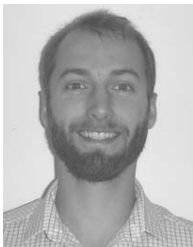
Stephen G. Matthews is a research assistant at the University of Bristol. He received a B.Sc. (Hons.) in computer science (2006), an M.Sc. in computational intelligence and robotics (2007) and a Ph.D. in computer science (2013) from De Montfort University, UK. His main research interests are evolutionary algorithms, fuzzy and genetic fuzzy systems, association rule mining, subspace clustering, temporal and intelligent data mining, and evolving systems.

Mario A. Gongora is a principal lecturer at the School of Computer Science and Informatics and the deputy director of the Centre for Computational Intelligence (CCI). He is a leading researcher in the area of applied CI, and for the last 10 years has been researching in the application of artificial intelligence technologies to complex problems and systems as well as intelligent data mining. He has extensive experience in systems integration for industry and process modelling using CI, and also valuable experience in exploiting IP from his research outcomes.

Adrian A. Hopgood is a pro vice-chancellor and the dean of the Sheffield Business School at Sheffield Hallam University. He has previously worked for De Montfort University, Nottingham Trent University, the Open University, Telstra Research Laboratories, and Systems Designers PLC. He has published over 100 papers and supervised 17 Ph.D. projects to completion. The third edition of his text book *Intelligent Systems for Engineers and Scientists* was published in 2012. He is a fellow of the British Computer Society. He holds a Ph.D. from the University of Oxford, an M.B.A. from the Open University, and a B.Sc. from the University of Bristol.