Multi-objective optimisation based fuzzy association rule mining method

Hui Zheng[1,2,3] · Jing He[1] · Qing Liu[4] · Jianhua Li[5] · Guangli Huang[6] · Peng Li[1,2]

Abstract

Fuzzy association rule mining (FARM) is a mainstream method to discover hidden patterns and association rules in quantitative data. It is essential to improve performance metrics, including quantity performance (e.g., the number of rules, the number of frequent itemsets) and quality performance (e.g., fuzzy support and confidence). The current approaches inadequately support optimisation of both quantity and quality performance. We propose a multi-objective optimisation algorithm for FARM (MOOFARM), where quantity and quality performance metrics are improved and validated simultaneously. The experimental evaluation conducted on a real dataset showcases the outstanding performance of MOOFARM against state-of-the-art works. In particular, at minimum support $=0.1$, minimum confidence $=0.7$, our MOOFARM increases the quantity performance up to 11 times. The proposed method improves the quality performance up to 71.05%.

Keywords Multi-objective functions · Optimisations · Fuzzy association rules

1 Introduction

A fuzzy association rule, such as {high blood pressure} $\xrightarrow{0.6}$ {having diabetes}, means that if a person suffers from high blood pressure, then there will be about a 60% probability of having diabetes. Fuzzy association rules contain fuzzy sets {having diabetes, 0.6}, which can express various types of vagueness and uncertainty in relations.

Fuzzy association rule mining, has been applied in Web Intelligence [1], since it can provide interpretable relations for quantitative data with fuzzy sets [2]. It has been widely applied

✉ Jing He
 lotusjing@gmail.com

✉ Peng Li
 lipeng@njupt.edu.cn

Extended author information available on the last page of the article

in web mining [3], recommendation systems [4], data analytics [5, 6], decision making [7, 8], natural language processing [9] and so on.

Enhancing metrics of fuzzy association rules' number is one of popular ways to improve the quality of fuzzy association rules. However, there usually are more than one metric for measuring fuzzy association rules, including quantity performance (e.g., the number of rules, the number of frequent itemsets) and quality performance (e.g., fuzzy support and confidence) [10], and each of them can be equally important. Thus, it is essential to optimise multiple objective functions for association rules with high quantity and quality discovery.

There are some existing reference works on the problem of multiple objective functions for association rule mining. Some papers apply stochastic optimisation methods, such as genetic optimisation (GA) algorithm, Ant Colony Optimisation (ACO) algorithm, cross-entropy (CE) algorithm, and reactive search optimisation (RSO) algorithm [11]. Some papers apply the deterministic optimisation method, such as Linear programming problems (LP), Mixed-integer linear programming problems (MILP), Non-linear programming problems (NLP), and Mixed-integer non-linear programming problems (MINLP) [12]. Some papers, such as [13], propose the simulated water-stream algorithm (SWA) by combining stochastic and deterministic optimisation. The SWA algorithm solves objective functions problems that are nondifferentiable.

However, there are still two challenges for existing work: 1) find high performance rules by selecting appropriate partitioning points for fuzzy sets with the given dataset; 2) optimise multiple metrics simultaneously to enhance the quality and quantity of fuzzy association rules.

Therefore, this paper proposes a multi-objective optimisation based fuzzy association rule mining (MOOFARM) method to optimise all metrics in every optimisation step. Then, the proposed method can generate fuzzy association rules with both high quality and quantity. Our contribution is threefold.

– The proposed multi-objective optimisation method, which optimises all objections simultaneously in each optimisation step (Section 3.2), can be applied in any other application.
– Our MOOFARM method improves both quantity and quality performance of fuzzy association rules by locating partitioning points for changeable fuzzy sets and division sets, as shown in Section 3.3.
– Experimental results also witness that our MOOFARM method outperforms its counterparts in terms of quality and quantity of fuzzy association rules, as shown in Section 4.

The rest of this paper is as follows. Section 2 shows existing fuzzy association rule mining methods with an optimisation process. We propose our multi-objective optimisation based fuzzy association rule mining (MOOFARM) method in Section 3. Following this, Section 4 illustrates our experiments and compares our proposed MOOFARM method with its counterparts. Finally, Section 5 concludes this paper.

2 Related works

The current fuzzy association rule mining methods for multiple optimisation include 1) a multi-objective approach; 2) a fuzzy logic-based approach; 3) a fuzzy-based multi-objective optimisation approach.

We first enumerate multi-objective optimisation works. Reference [14] applies a multi-objective genetic approach that can find association rules and improve the accuracy of classification. The fuzzy system generates a classier with high performance in a real-life power system. Reference [15] inaugurates the concept of fuzzy association rule mining and proposes a multi-object genetic algorithm (GA). The authors demonstrate the improvement of strongness, interestingness and comprehensibility. Reference [16] uses a multi-objective genetic approach to extract and optimise fuzzy rules. A computer activity dataset tests the effectiveness.

In addition, more efforts have been made to the multi-objective optimisation. The authors apply the bat algorithm and introduce a cooperative master-slave strategy for different subpopulations in [17]. Various experiments illustrate the bat algorithm outperforming other bio-inspired ones thanks to the association rule mining method. Furthermore, the same authors improve the bat algorithm based association rule mining method in [18]. Their study on cooperative strategies between swarms results in a multi-swarm optimisation method. Reference [17] also applies a hybrid topology by merging the ring and the master-slave methods in [17]. The results of multi-swarm optimisation show a clear superiority against its counterparts in terms of time and rule quality. Reference [19] discusses and compares two genetic algorithms: 1) Genetic Cooperative-Competitive Learning algorithm (GFS.GCCL algorithm); 2) Structural learning algorithm (SLAVE algorithm). The authors employ two public datasets: the Iris dataset and the Wine dataset in the experiments. The experimental results show that the SLAVE algorithm performs better than the GFS.GCCL algorithm in these two datasets. Reference [20] demonstrates the dimension-reduction technique in transactions and itemsets. The authors apply fuzzy and whale optimisation for frequent itemsets discovery and association rule mining. The proposed algorithm is comparable to particle swarm optimisation genetic algorithm and frequent fuzzy itemset-miner, as illustrated in the experimental results.

Next, we discuss fuzzy logic-based optimisation. Reference [21] applies to learn automata (LA) representation with reinforcement-based optimisation method (LA-OMF) to discover fuzzy sets of the time spent by users on each web page (TMFs) when mining fuzzy association rules. The experimental evaluation on a real dataset shows that the LA-OMF method is more efficient in discovering optimised membership function-based fuzzy association rules. Ai et al. [22] presents an optimisation method to optimise partitioning points for fuzzy sets in the fuzzy association rule mining process adaptively. The proposed strategy can predict behaviours with unknown system dynamics. The authors implement an FNN-based soft-sensor to adjust the number of fuzzy sets and optimal image features for better control performance. The strategy is validated in simulations and real-life experiments.

Then, there is also some work on fuzzy-based multi-objective optimisation approach. Reference [23] proposes a multi-objective optimisation method and applies it to fuzzy association rule mining to improve rules' quality and quantity compared with the FARM method proposed in [24]. The proposed optimised method can reach a better result for at least one of the objectives in each optimisation step. The authors innovate the previous strategy and achieve higher quality and quantity for fuzzy association rules in [25], inspiring us to refine the procedures for multiple objectives optimisation. However, it remains open to obtaining better results for multiple objectives in each optimisation step of association rule mining.

3 Multiple objective optimisation for association rule performance metrics

The multi-objective optimisation method based on gradient descent and Schmidt orthogonal projection mainly includes the following two key steps: 1) Clarify the rule quality standard that needs to be optimised, that is, the objective function in multi-objective optimisation; 2) Use Schmidt orthogonalised projection and its matrix decomposition obtain the direction vector that can optimise multiple objective functions at the same time, and prove through theorem that the multi-objective optimisation algorithm can optimise all objective functions for each the iterative process.

Note that the proposed multi-objective optimisation method can solve the general problem of optimising multiple objectives. The objective functions relating to association rules are just one of its applications.

3.1 Association rules of interestingness and objective functions

Before discussing our objective functions, we borrow the concept of interestingness proposed by Silberschatz and Tuzhilin [26]. The definition of interestingness is shown in Definition 1.

Definition 1 Interestingness in association rule mining describes that an object is valuable on a given occasion. It usually explains that a condition of an association rule or an item is unexpected or actionable by the user in a fixed data set or on a given occasion.

To measure the 'interestingness' of an association rule r, we apply the following function from our previous work [23] in multiple objective optimisation processes.

$$\varphi(r) = \min \left\{ \begin{array}{l} \text{Supp}(r) - \text{minSupp} \\ \text{Conf}(r) - \text{minConf} \\ \text{CF}(r) - \text{minCF} \end{array} \right\}. \tag{1}$$

This formula (1) is applied to control the performance of mined association rules in this paper, where minSupp, minConf and minCF present corresponding user-defined threshold, respectively. Then, the value of function $\varphi(r)$ larger means the mined association rules' quality is better. To further optimise the quality of mined rules and compare them with our previous work, we also apply the following $\varphi(r)$ based objective functions in [23].

- $\Phi_1 = \max\limits_{r} \varphi(r)$: the measurement of the best quality rules among all mined fuzzy association rules;
- $\Phi_3 = \max\limits_{r_1, r_2, r_3} \sum_{i=1}^{3} \varphi(r_i)$: the average measurement of the three best quality rules among all mined fuzzy association rules;
- Φ_5: the average measurement of the five best quality rules among all mined fuzzy association rules;

- Φ_{10}: the average measurement of the ten best quality rules among all mined fuzzy association rules;
- $\Phi_{n/2}$: the average measurement of the half best quality rules among all mined fuzzy association rules;

We will use the worst quality rules to fill the deficient ones in calculating these objective functions in our experiments. The performance of these objectives will also be compared as measurements for the quality of fuzzy association rules.

3.2 Construction of our multiple objective optimisations

How to optimise multiple objective functions at the same time is a primary problem in the process of designing multiple objective optimisation algorithms. The basic idea is to construct a matrix according to the gradient of all objective functions and perform Schmidt orthogonal decomposition of the constructed matrix. Then, our optimisation process finds the optimal iterative direction vector of the multiple objective functions according to Schmidt orthogonal decomposition.

Suppose there are m objective functions, then they are denoted as

$$\varphi_1, \varphi_2, ..., \varphi_m$$

respectively. The gradient of these m objective functions are denoted as

$$\nabla\varphi_1, \nabla\varphi_2, ..., \nabla\varphi_m.$$

All gradients of the objective functions form a $n \times m$ matrix G. Then, we have

$$G = [\nabla\varphi_1, \nabla\varphi_2, ..., \nabla\varphi_m].$$

As matrix G has n rows, which means the matrix G^T has n columns.

To find a n dimension vector $\mathbf{x}$, so that

$$G^T\mathbf{x} = \mathbf{y} = [y_1, y_2, ..., y_m]^T,$$

satisfies condition $y_1 > 0, y_2 > 0, ..., y_m > 0$, then, we can obtain n dimension vector $\mathbf{x}$. This gradient vector $\mathbf{x}$ is the solution that can optimise m objective functions simultaneously. To construct this gradient vector $\mathbf{x}$, we define the following Lemmas [27, 28].

Definition 2 (Vector origin perpendicular) In n dimension linear space, for $\forall$ non-zero n dimension vector $\mathbf{v}$, the vectors which perpendicular to the vector $\mathbf{v}$ and cross the origin point's hyperplane are denoted as the origin perpendicular [29]. They are shown as follows: $\perp_{\mathbf{v}}$.

Definition 3 (Vector group of vertical subspace) In n dimension linear space, $\forall k, k < n$ non-zero n dimension vector $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$, the vector group of vertical subspace [28] is shown as follows:

$$\perp_{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k} := \perp_{\mathbf{v}_1} \cap \perp_{\mathbf{v}_2} \cap ... \cap \perp_{\mathbf{v}_k}.$$

Definition 4 (Vertical projection of vector on vector group) In the n dimensional linear space, for any given $k < n$ non-zero n dimensional vector $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ and another n

dimensional vector $\mathbf{x}$, vector $\mathbf{x}$ about the vector group $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is defined as $\mathbf{x}$ in projection $\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k}$ [27],

$$\mathcal{P}_{\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k}} \mathbf{x}.$$

Lemma 1 (Calculation method of vector vertical projection on vector group) In a n dimensional linear space, let $\{k : k < n\}$ non-zero n dimensional vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ and another n dimensional vector $\mathbf{x}$ form a linearly independent vector group of $k + 1$ vectors, then Schmidt orthogonalisation can get matrix decomposition

$$[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{x}] = QU,$$

where

$$Q = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{k+1}]$$

is a $n \times (k + 1)$ matrix and

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{2}$$

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,k} & u_{1,k+1} \\ 0 & u_{2,2} & \cdots & u_{2,k} & u_{2,k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{k,k} & u_{k,k+1} \\ 0 & 0 & \cdots & 0 & u_{k+1,k+1} \end{bmatrix}$$

is an upper triangular matrix with all positive diagonal elements and has the following equation

$$\mathcal{P}_{\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k}} \mathbf{x} = u_{k+1,k+1} \mathbf{q}_{k+1}.$$

Proof First, let's see the Schmidt Orthogonalisation process, as shown below:

$$u_{1,1} = \sqrt{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle},$$

$$\mathbf{q}_1 = \frac{1}{u_{1,1}} \mathbf{v}_1,$$

$$u_{1,2} = \langle \mathbf{q}_1, \mathbf{v}_2 \rangle,$$

$$\tilde{\mathbf{v}}_2 = \mathbf{v}_2 - u_{1,2} \mathbf{q}_1,$$

$$u_{2,2} = \sqrt{\langle \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_2 \rangle},$$

$$\mathbf{q}_2 = \frac{1}{u_{2,2}} \tilde{\mathbf{v}}_2,$$

$$\vdots$$

$$u_{i,j} = \left\langle \mathbf{q}_i, \mathbf{v}_j \right\rangle \quad \forall i = 1, \ldots, j-1$$

$$\tilde{\mathbf{v}}_j = \mathbf{v}_j - \sum_{i=1}^{j-1} u_{i,j} \mathbf{q}_i,$$

$$u_{j,j} = \sqrt{\left\langle \tilde{\mathbf{v}}_j, \tilde{\mathbf{v}}_j \right\rangle},$$

$$\mathbf{q}_j = \frac{1}{u_{j,j}} \tilde{\mathbf{v}}_j,$$

$$\vdots$$

$$u_{i,k+1} = \left\langle \mathbf{q}_i, \mathbf{x} \right\rangle \quad \forall i = 1, \ldots, k$$

$$\tilde{\mathbf{x}} = \mathbf{x} - \sum_{i=1}^{k} u_{i,k+1} \mathbf{q}_i,$$

$$u_{k+1,k+1} = \sqrt{\left\langle \tilde{\mathbf{x}}, \tilde{\mathbf{x}} \right\rangle},$$

$$\mathbf{q}_{k+1} = \frac{1}{u_{k+1,k+1}} \tilde{\mathbf{x}},$$

The linear independence condition between the vector group $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{x}$ is for all i = 1,2,...,$k+1$ both ensure that $u_{i,i} > 0$, so that when calculating $\mathbf{q}_i$, $u_{i,i}$ can be used as the denominator.

From the above Schmidt orthogonalisation process, we can also know that there are the following relations:

$$\mathbf{v}_j \in \mathrm{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_j\} \quad \forall j = 1, 2, \ldots, k$$

Apart from the above formula, it has orthogonalisation relations as follows.

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Therefore,

$$\mathbf{q}_{k+1} \perp \mathbf{v}_j \quad \forall j = 1, 2, \ldots, k,$$

In other words,

$$\mathbf{q}_{k+1} \in \perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k},$$

From $q_1, q_2, \ldots, q_k$ can be expressed linearly by $v_1, v_2, \ldots, v_k$, we have

$$\mathbf{q}_j \perp\!\!\!\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k},$$

where $j < k + 1$, then,

$$\sum_{i=1}^{k} u_{i,k} \mathbf{q}_i \perp\!\!\!\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k}$$

Finally, with

$$\mathbf{x} = \sum_{i=1}^{k+1} u_{i,k+1} \mathbf{q}_i,$$

$$\mathbf{q}_{k+1} \in \perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k},$$

and

$$\sum_{i=1}^{k} u_{i,k} \mathbf{q}_i \perp\!\!\!\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k},$$

we have

$$\mathcal{P}_{\perp_{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k}} \mathbf{x} = u_{k+1,k+1} \mathbf{q}_{k+1}.$$

□

Lemma 2 Let $m \leq n$ and $m \times n$ real matrix A full row rank, That is, the rank of the matrix $Ar(A) = m$, the non-zero vector $\mathbf{x}$ can be found so that

$$A\mathbf{x} = \mathbf{y} = (y_1, y_2, \ldots, y_m)^T, \tag{3}$$

satisfy $y_1 \geq 0, y_2 \geq 0, \ldots, y_m \geq 0$, and at least one of them makes $y_i > 0$ true.

Proof Denote A's i-th row is $\mathbf{a}_{i,*}$, $\forall$ integer in $1, 2, \ldots, m$, and having

$$\mathbf{x} = \mathcal{P}_{\perp_{\mathbf{a}_1,*, \ldots, \mathbf{a}_{i-1,*}, \mathbf{a}_{i+1,*}, \ldots, \mathbf{a}_{m,*}}} \mathbf{a}_{i,*},$$

Then, with Lemma 1, we can ensure that $\mathbf{x}$ is non-zero, and the portion vector $y_1, y_2, \ldots, y_m$ of

$$\mathbf{y} = A\mathbf{x}$$

has

$$y_j = \begin{cases} \langle x, x \rangle, & i = j \\ 0, & i \neq j \end{cases}$$

□

Theorem 1 Suppose $m \leq n$ and $m \times n$ real matrix A is full row rank, then

$$\mathbf{x} := \sum_{i=1}^{m} \mathcal{P}_{\perp_{\mathbf{a}_1,*, \ldots, \mathbf{a}_{i-1,*}, \mathbf{a}_{i+1,*}, \ldots, \mathbf{a}_{m,*}}} \mathbf{a}_{i,*} \tag{4}$$

For portion vector $y_1, y_2, \ldots, y_m$ of

$$\mathbf{y} = A\mathbf{x},$$

which satisfies

$$y_1 > 0, y_2 > 0, \ldots, y_m > 0$$

Proof According to the proof processing of Lemma 2, for

$$\mathbf{x}_i := \mathcal{P}_{\perp_{\mathbf{a}_{1,*},\ldots,\mathbf{a}_{i-1,*},\mathbf{a}_{i+1,*},\ldots,\mathbf{a}_{m,*}}} \mathbf{a}_{i,*}$$

vector $A\mathbf{x}_i$, only the i-th portion equals $\langle \mathbf{x}_i, \mathbf{x}_i \rangle > 0$, while other portions are set to 0. Then, we have the sum of them,

$$\mathbf{x} = \sum_{i=1}^{m} \mathbf{x}_i$$

satisfies the conclusion of this theorem. □

Note that the Theorem 1 not only proves the existence of $\mathbf{x}$, but also gives the construction method of $\mathbf{x}$. The solution method is easy to realise with a computer program. Based on this theorem, the process of determining the search direction for multi-objective optimisation is the construction of the gradient construction matrix corresponding to the objective function.

$$A = \begin{bmatrix} \nabla\varphi_1^T \\ \nabla\varphi_2^T \\ \vdots \\ \nabla\varphi_m^T \end{bmatrix}, \tag{5}$$

Then, we can solve the corresponding $\mathbf{x}$ in the formula (4) via Theorem 1, that is, $A\mathbf{x} = \mathbf{y}$ $= (y_1, y_2, \ldots, y_m)^T$ satisfy the relationship $y_1 > 0, y_2 > 0, \ldots, y_m > 0$. The vector $\mathbf{y}$ in Theorem 1 corresponds to the degree of optimisation of the multi-objective function. The direction vector $\mathbf{x}$ that satisfies the theorem can be solved simultaneously. The direction vector for optimising the multi-objective function is, therefore, solved.

Although the requirement of the theorem on the rank of $Ar(A) = m$ increases the limitation of the theorem. In practical applications, the probability of $r(A) = m$ in the iterative optimisation process is high. Theoretically, the optimal solution of the optimisation problem will degenerate the matrix A to $r(A) = 1$. However, the optimal solution is not attainable in a finite step, and the detection of the optimal solution is comparative. It is easier in solving $\mathbf{x}$ when the corresponding row vector of A is not full rank.

Algorithm 1 Find the direction for multi-objective optimisation.

Input : Multi-objective function object-rule quality evaluation index: $\varphi_1, \varphi_2, \ldots, \varphi_m$, where

$$\varphi(r) = \min \left\{ \begin{array}{l} \text{Supp}(r) - \text{minSupp} \\ \text{Conf}(r) - \text{minConf} \\ \text{CF}(r) - \text{minCF} \end{array} \right\},$$

$\varphi_1, \varphi_2, \ldots, \varphi_m$ in equation (1).

Output: direction vector: η - All objective functions keep increasing along this direction.

1 Initialisation: Construct a matrix based on a multi-objective function object, where $\nabla\varphi_1^T, \nabla\varphi_2^T, \cdots, \nabla\varphi_m^T$ are the gradients corresponding to the objective function object. This paper approximates the gradient value for the partitioning points of fuzzy sets according to Richard's extrapolation method.

$$A = \begin{bmatrix} \nabla\varphi_1^T \\ \nabla\varphi_2^T \\ \vdots \\ \nabla\varphi_m^T \end{bmatrix}$$

2 Calculate the corresponding direction about the matrix A by the Theorem 1

$$\mathbf{x} = \sum_{i=1}^{m} \mathcal{P}_{\perp_{\mathbf{a}_{1,*},\ldots,\mathbf{a}_{i-1,*},\mathbf{a}_{i+1,*},\ldots,\mathbf{a}_{m,*}}} \mathbf{a}_{i,*};$$

3 **return** $\eta = \mathbf{x}$;

From Algorithm 1, we can see that the process of finding the direction for multi-objective optimisation, as shown in lines $1 - 3$. Line 1 shows the initialisation step. Line 2 calculate the direction according to Theorem 1. Then, line 3 returns the direction vector η.

Theorem 2 In each optimisation iteration process of multi-objective optimisation, if the vector group formed by the gradient vectors of these objective functions is linearly independent, there is a direction vector η that can optimise all objective functions at the same time.

Proof It is a direct inference of the Theorem 1. Note that the linearity of the vector group here is only to show that the row of the vector group is full of rank. □

Combined with the Algorithm 1, the Theorem 2 is further explained as follows:

- The selected interval division points can be gradually optimised through the iterative optimisation target process in Algorithm 1.
- Fuzzy set and its membership value have no corresponding changes in Algorithm 1, and the update process needs further research.

Determining the objective function of the multi-objective optimisation algorithm is the process of continuously optimising the quality evaluation metrics. This process is the core

of the multi-objective optimisation algorithm and the process of discovering better association rules. Therefore, while ensuring the current multi-objective optimisation process, it is necessary to consider the construction process of fuzzy sets and related membership functions.

3.3 Find the direction for global multi-objective optimisation

This subsection further improves our multi-objective optimisation method by applying a random partitioning point for division sets. Also, we compare the optimisation parameters of our MOOFARM approach with that of our previous work, DOFARM.

Algorithm 2 Find the direction for global multi-objective optimisation.

    Input   : Multi-objective function object-rule quality evaluation index: $\varphi_1, \varphi_2, \ldots, \varphi_m$, where

$$\varphi(r) = \min \left\{ \begin{array}{l} \text{Supp}(r) - \text{minSupp} \\ \text{Conf}(r) - \text{minConf} \\ \text{CF}(r) - \text{minCF} \end{array} \right\},$$

          $\varphi_1, \varphi_2, \ldots, \varphi_m$ in equation (1).

    Output: global optimised direction vector: η (Make sure every objective function keeps an increasing trend along this direction).

1 Initialisation Step:

2 Construct a matrix based on a multi-objective function object, where $\nabla \varphi_1^T, \nabla \varphi_2^T, \cdots,$ $\nabla \varphi_m^T$ are the gradients corresponding to the objective function object. This paper approximates the gradient value according to Richard's extrapolation method.

$$A = \begin{bmatrix} \nabla \varphi_1^T \\ \nabla \varphi_2^T \\ \vdots \\ \nabla \varphi_m^T \end{bmatrix}$$

3 Initialise current best parameters bestPara = the original values of parameters $\mathbf{x}$;

4 Initialise current best objective values bestObj = $A * \mathbf{x}$;

5 for(int i = 0; i < user-defined numbers; i++)

6 {

7   Randomly choose partitioning points for division sets for multi-objective optimisation;

8     $\mathbf{x}$ = find the direction for multi-objective optimisation according to Algorithm 1;

9   if($A * \mathbf{x}$ > bestObj)

10   {

11   bestPara = $\mathbf{x}$;

12   bestObj = $A * \mathbf{x}$;

13   }

14 }

15 **return** $\eta = bestPara$;

According to Algorithm 2, we can see our global multi-objective optimisation processes from lines 1 − 15. Lines 1 − 4 are our initialisation steps. Then, we apply our randomly obtained parameters for multi-objectives and apply our proposed optimisation method by

lines 5 − 14. Next, we compare the best objectives for each group of randomly obtained parameters by line 7 and keep the best Parameters that can achieve the best objectives according to Algorithm 1 in line 8. In this way, we can find the best direction for our global multi-objective optimisation with lines 11 − 13. We return the best direction in line 15.

According to our previous work [25], division sets construct the fuzzy sets for fuzzy association rules, and five division sets can generate three fuzzy sets. As shown in Figure 1, there are five division sets, and the two partitioning points for the three fuzzy sets are related to the second division set and the fourth division set separately. Therefore, two partitioning points of fuzzy sets are adjustable by changing four partitioning points of division sets.

The most related work DOFARM method adjusts and optimises the partitioning points of fuzzy sets to achieve high performance of fuzzy association rules. In contrast, our proposed MOOFARM optimises the partitioning points for both fuzzy sets and division sets. In this way, our proposed method MOOFARM will be more robust and can achieve better optimisation results. Then, we can mine higher quality and quantity of fuzzy association rules.

4 Experimental results

This section describes our experimental study of the proposed method. We also compare the proposed method with the most related work to show the excellent performance of our proposed multi-objective optimisation based fuzzy association rule mining (MOOFARM) method. We also compare our proposed MOOFARM method with a general fuzzy association rule mining (FARM) method [24] and our previous work DOFARM method [25]. Dataset 'pima-indians-diabetes' [30] from UCI Machine Learning Data is applied in our experiments.

4.1 Comparison with the number of frequent itemsets and fuzzy association rules

This subsection compares our MOOFARM method with existing methods: the FARM method and DOFARM method with the number of frequent itemsets and fuzzy association rules.
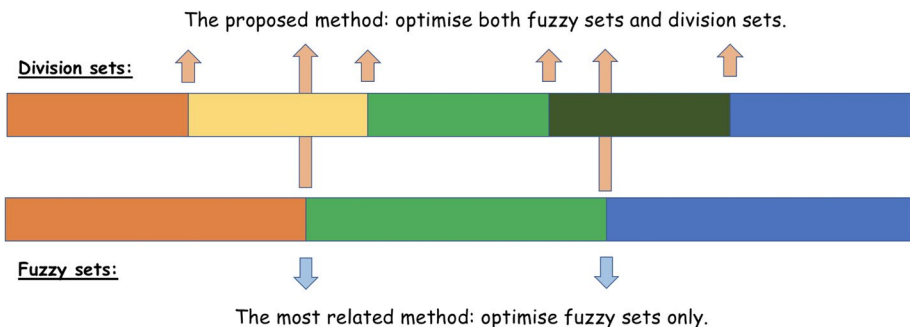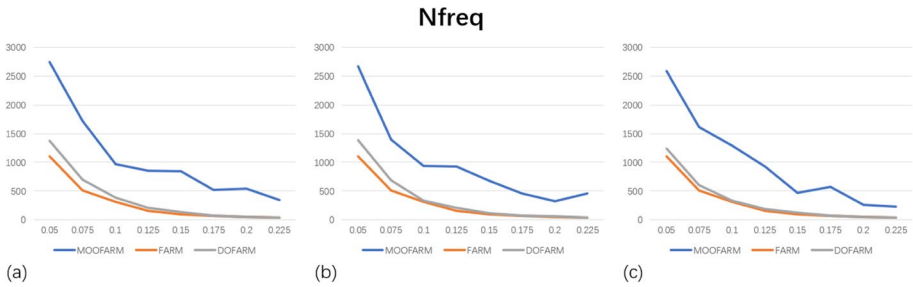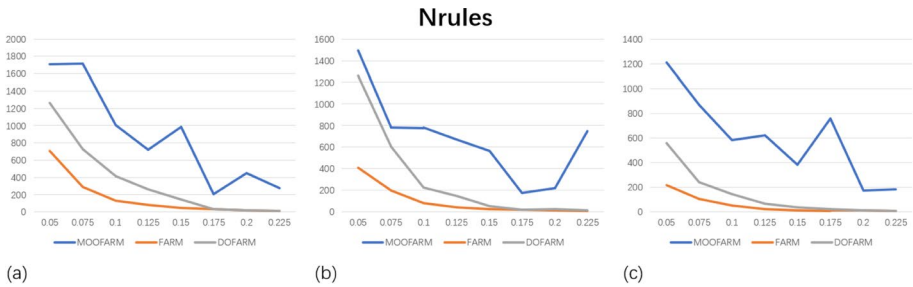

Fig. 1 Fuzzy set optimisation comparison of our proposed method and the most related work

Fig. 2 Comparison of the number of frequent item sets. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7

Fig. 3 Comparison of the number of fuzzy association rules. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7

Figure 2 illustrates the number of frequent itemsets (Nfreq) among our proposed MOO-FARM method, DOFARM method, and FARM method. The performance of three different minimum confidences: 0.5,0.6,0.7 are shown in sub-figures (a), (b), (c) in Figure 2 respectively. Our proposed MOOFARM method has a larger number of frequent itemsets than the numbers of both the FARM and DOFARM methods, such as at minimum support $=0.1$, minimum confidence $=0.7$, our MOOFARM (≈ 1300) is higher than both of DOFARM (≈ 400) and FARM (≈ 400) by $(1300 - 400)/400 = 2.25 = 225\%$ (2.25) times.

Figure 3 compares the number of fuzzy association rules (Nrules) among our proposed MOOFARM method, DOFARM method, and FARM method. For each different minimum confidence in: 0.5,0.6,0.7, our proposed MOOFARM method has a larger number of fuzzy association rules than the numbers of both the FARM method and DOFARM method. Similarly, taking minimum support $=0.1$, minimum confidence $=0.7$ as an example, our MOOFARM (≈ 600) is higher than both of DOFARM (≈ 100) and FARM (≈ 50) by 5 times and 11 times separately.

4.2 Comparison with the quality of fuzzy association rules

This subsection compares our MOOFARM method with existing methods: the FARM method and DOFARM method the quality of fuzzy association rules. The quality is presented in this paper according to $\varphi(r)$ as shown in formula (1).

Figure 4 compares the quality of fuzzy association rules (Best): $\varphi(r)$ as shown in formula (1) among our proposed MOOFARM method, DOFARM method, and FARM
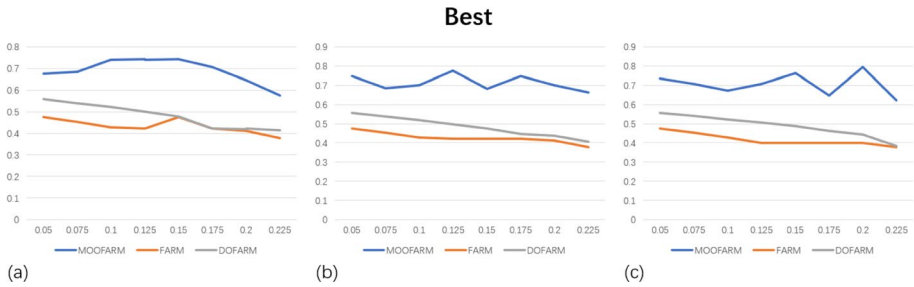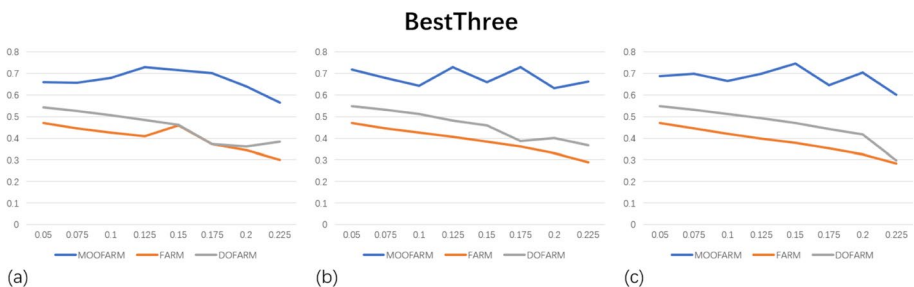
Fig. 4 Comparison of the best fuzzy association rule. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7

method. Selecting different minimum confidence: 0.5,0.6,0.7 in Figure 4, our proposed MOOFARM method always has better quality $\varphi(r)$ of fuzzy association rules than that of both the FARM method and DOFARM method. For example, at minimum support $=0.1$, minimum confidence $=0.7$, our MOOFARM (≈ 0.68) is higher than DOFARM (≈ 0.52) and FARM (≈ 0.42) by 30.77 % and 61.90 %.

As shown in Figure 5, the quality $\varphi(r)$ of fuzzy association rules (BestThree) of our proposed MOOFARM method is better than $\varphi(r)$ of DOFARM and FARM method at different minimum confidences: 0.5,0.6,0.7. At minimum support $=0.1$, minimum confidence $=0.7$, for example, our MOOFARM (≈ 0.67) is higher than DOFARM (≈ 0.51) and FARM (≈ 0.42) by 31.37 % and 59.52 %.

Figure 6 shows that the quality $\varphi(r)$ of fuzzy association rules (BestFive) of our proposed MOOFARM method is higher than that of DOFARM and FARM method at different minimum confidences: 0.5,0.6,0.7. For instance, when minimum support $=0.1$, minimum confidence $=0.7$, our MOOFARM (≈ 0.66) is higher than DOFARM (≈ 0.5) and FARM (≈ 0.4) by 32 % and 65 %.
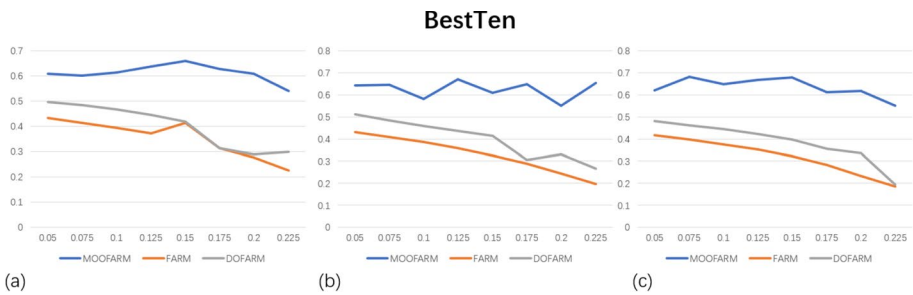
Figure 7 illustrates the quality of fuzzy association rules (BestTen) among our proposed MOOFARM method, DOFARM method, and FARM method at different minimum confidences: 0.5,0.6,0.7. Our proposed MOOFARM method's quality $\varphi(r)$ of fuzzy association rules is always higher than that of the FARM and DOFARM. Taking minimum support $=0.1$, minimum confidence $=0.7$ as an example, our MOOFARM (≈ 0.65) is higher than DOFARM (≈ 0.45) and FARM (≈ 0.38) by 44.44 % and 71.05 %.
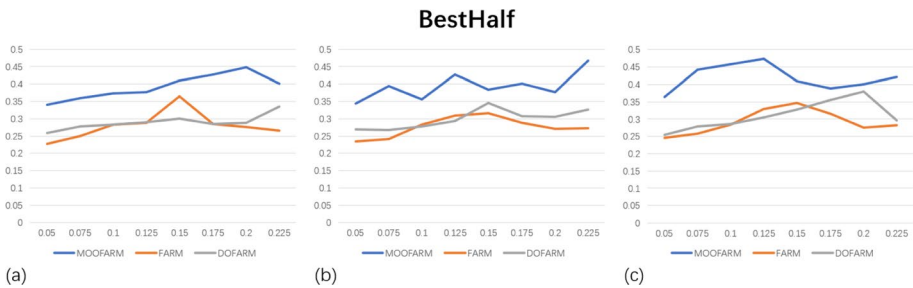

Fig. 5 Comparison in the best three fuzzy association rules. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7
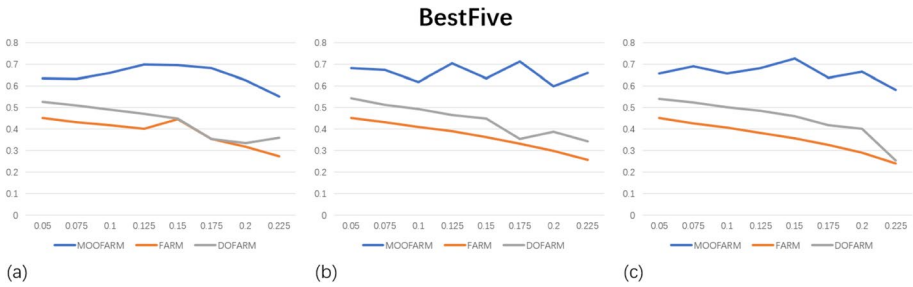
Fig. 6 Comparison in the best five fuzzy association rules. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7

Fig. 7 Comparison in the best ten fuzzy association rules. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7

Fig. 8 Comparison in the best half fuzzy association rules. **a** Confidence=0.5; **b** Confidence=0.6; **c** Confidence=0.7

Figure 8 applies different minimum confidence: 0.5,0.6,0.7 to better present the quality performance of all three methods. Our proposed MOOFARM method's quality $\varphi(r)$ of fuzzy association rules are much higher than that of both the FARM method and DOFARM method. For example, at minimum support $=0.1$, minimum confidence $=0.7$, the proposed MOOFARM method (≈ 0.46) increases the quality of fuzzy association rules (BestHalf) by 64.28 % compared to both DOFARM method (≈ 0.28), and FARM method (≈ 0.28).

According to Figures 2 and 3, our proposed MOOFARM method increases the number of frequent item sets and the number of fuzzy association rules than the FARM method and DOFARM method. Also, in Figures 4, 5, 6, 7 and 8, it is evident that the quality of fuzzy

association rules mined by our proposed MOOFARM method is higher than that of its two counterparts. Specifically, at minimum support $=0.1$, minimum confidence $=0.7$, the proposed MOOFARM method's quantity of fuzzy association rules increases approximately $2.25 - 5$ times (DOFARM) and $2.25 - 11$ times; while the quality of MOOFARM method has been improved around $30.77 - 62.28\%$ and $59.52 - 71.05\%$ for DOFARM and FARM respectively.

5 Conclusion

This paper first proposes a general multi-objective optimisation method to optimise all objections for each optimisation step simultaneously. Then, it is applied in fuzzy association rule mining to build our multi-objective optimisation-based fuzzy association rule mining (MOOFARM) model. In addition, we employ metrics in optimisation to improve both association rule's quality and quantity performance. Moreover, our MOOFARM method achieves better performance as it optimises partitioning points in two rounds: different fuzzy sets and dynamic division sets. Experiments also show that our MOOFARM method outperforms its two counterparts. Specifically, at minimum support $=0.1$, minimum confidence $=0.7$, our MOOFARM increases the quantity metrics up to 11 times; for the quality of fuzzy association rules, the growth of the proposed MOOFARM method is up to 71.05%.

References

1. Kalia, H., Dehuri, S., Ghosh, A.: A survey on fuzzy association rule mining. International Journal of Data Warehousing and Mining (IJDWM) **9**(1), 1–27 (2013)
2. Alcala-Fdez, J., Alcala, R., Herrera, F.: A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. IEEE Trans. Fuzzy Syst. **19**(5), 857–872 (2011)
3. Arotaritei, D., Mitra, S.: Web mining: a survey in the fuzzy framework. Fuzzy Set. Syst. **148**(1), 5–19 (2004)
4. Wanaskar, U., Vij, S., Mukhopadhyay, D.: A hybrid web recommendation system based on the improved association rule mining algorithm. arXiv:1311.7204 (2013)
5. Jindal, A., Dua, A., Kumar, N., Das, A.K., Vasilakos, A.V, Rodrigues, J.J.P.C.: Providing healthcare-as-a-service using fuzzy rule based big data analytics in cloud computing. IEEE Journal of Biomedical and Health Informatics **22** (5), 1605–1618 (2018)
6. Pérez-Alonso, A., Blanco, I.J., Serrano, J.M., González-González, L.M.: Incremental maintenance of discovered fuzzy association rules. Fuzzy Optim. Decis. Making, 1–21 (2021)
7. Noor, N.M.M., Wan, M.F., Nawawi, W., Ghazali, A.F.: Supporting decision making in situational crime prevention using fuzzy association rule. In: 2013 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), pp. 225–229. IEEE (2013)
8. Shang, H., Duan, L u, Zhou, Q.: Early warning of enterprise finance risk of big data mining in internet of things based on fuzzy association rules. Neural Comput. & Applic. **33**(9), 3901–3909 (2021)
9. Cambria, E., White, B.: Jumping nlp curves: A review of natural language processing research. IEEE Computational Intelligence Magazine **9**(2), 48–57 (2014)

10. Muyeba, M., Sulaiman Khan, M., Coenen, F.: Fuzzy weighted association rule mining with weighted support and confidence framework. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp 49–61. Springer (2008)

11. Najim, K., Ikonen, E., Daoud, A.-K.: Stochastic processes: estimation, optimisation and analysis. Elsevier (2004)

12. Lin, M.-H., Tsai, J.-F., Yu, C.-S.: A review of deterministic optimization methods in engineering and management. Math. Probl. Eng. 2012 (2012)

13. Cheung, Y.-M., Gu, F, Liu, H.-L., Tan, K.C., Huang, H.: Objective-domain dual decomposition: An effective approach to optimizing partially differentiable objective functions. IEEE Transactions on Cybernetics **50**(3), 923–934 (2018)

14. Li, F., Jiaju, Q.: Short-term load forecasting for anomalous days based on fuzzy multi-objective genetic optimization algorithm. Proceedings of the CSEE **25**(10), 29–34 (2005)

15. Kaya, M.: Multi-objective genetic algorithm based approaches for mining optimized fuzzy association rules. Soft Computing **10**(7), 578–586 (2006)

16. Santhi Thilagam, P., Ananthanarayana, V.S.: Extraction and optimization of fuzzy association rules using multi-objective genetic algorithm. Pattern. Anal. Applic. **11**(2), 159–168 (2008)

17. Heraguemi, K.E., Kamel, N., Drias, H.: Multi-population cooperative bat algorithm for association rule mining. In: Computational Collective Intelligence, pp 265–274. Springer (2015)

18. Heraguemi, K.E., Kamel, N., Drias, H.: Multi-swarm bat algorithm for association rule mining using multiple cooperative strategies. Appl. Intell. **45**(4), 1021–1033 (2016)

19. Kar, S., Kabir, M.M.J.: Comparative analysis of mining fuzzy association rule using genetic algorithm. In: 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp 1–5. IEEE (2019)

20. Sharmila, S., Vijayarani, S.: Association rule mining using fuzzy logic and whale optimization algorithm. Soft. Comput. **25**(2), 1431–1446 (2021)

21. Anari, Z., Hatamlou, A., Anari, B., Masdari, M.: Optimizing membership functions using learning automata for fuzzy association rule mining. Journal of AI and Data Mining **8**(4), 491–514 (2020)

22. Ai, M., Xie, Y., Xie, S., Zhang, J., Gui, W.: Fuzzy association rule-based set-point adaptive optimization and control for the flotation process. Neural Comput. & Applic. **32**(17), 14019–14029 (2020)

23. Zheng, H., He, J., Huang, G., Zhang, Y.: Optimized fuzzy association rule mining for quantitative data. In: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp 396–403. IEEE (2014)

24. Delgado, M., Marin, N., Sanchez, D., Vila, M. -A.: Fuzzy association rules: general model and applications. IEEE Trans. Fuzzy Syst. **11**(2), 214–225 (2003)

25. Zheng, H., He, J., Huang, G., Zhang, Y., Wang, H.: Dynamic optimisation based fuzzy association rule mining method. Int. J. Mach. Learn. Cybern. **10**(8), 2187–2198 (2019)

26. Silberschatz, A., Tuzhilin, A.: What makes patterns interesting in knowledge discovery systems. IEEE Transactions on Knowledge and Data Engineering **8**(6), 970–974 (1996)

27. Wilkinson, J.H., Bauer, F.L., Reinsch, C.: Linear Algebra, vol. 2. Springer, Berlin (2013)

28. Demmel, J.W.: Applied numerical linear algebra. SIAM (1997)

29. Strang, G., Strang, G., Strang, G., Strang, G.: Introduction to Linear Algebra, vol. 3. Press Wellesley, Wellesley-Cambridge (1993)

30. Rossi, R.A., Ahmed, N.K.: The network data repository with interactive graph analytics and visualization. In: AAAI (2015)

Authors and Affiliations

Hui Zheng[1,2,3] [iD] **· Jing He[1] · Qing Liu[4] · Jianhua Li[5] · Guangli Huang[6] · Peng Li[1,2]**

Hui Zheng
hui.zheng@aliyun.com

Qing Liu
q.liu@data61.csiro.au

Jianhua Li
jack.li@deakin.edu.au

Guangli Huang
belinda.huang@deakin.edu.au

[1] School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China

[2] Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Jiangsu Province, Nanjing, China

[3] Swinburne University and Technology, Melbourne, Australia

[4] Software and Computational Systems Program, Data 61, CSIRO, Canberra, Australia

[5] School of Info Technology, Deakin University Burwood Campus, Melbourne, Australia

[6] School of Information Technology, Deakin University, Geelong, Australia