International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 18 (2015) pp 38643-38651


# An Optimal Weighted Fuzzy Rule Mining Classification On Multidimension Data Stream

**P.Velvadivu**
Research Scholar,
Department of Computing,
Coimbatore Institute of Technology
Coimbatore
velvadivu@gmail.com

**Dr.C.Duraisamy,**
Dean,
School of Science and Humanities,
Kongu Engineering College
Perundurai
cdssh@gmail.com

## ABSTRACT
Classification has been used for designing and structuring any types of datasets and applications including graphs, networks to medical diagnosis and keyword extraction and so on. Many research works were conducted for efficient classification using data mining techniques like Swift Rule and Piecewise Polynomial Modeling. However, there is a lack of study for effective classification on filtered data stream using fuzzy classification and also pattern based classification becomes a critical issue. In this work, an effective framework called Optimal Weighted Fuzzy Rule Mining Classification (OW-FRMC) on filtered data stream is designed. The Weighted Fuzzy rule permits to classify the multidimensional data with bipolar property and therefore improves the classification ratio. The bipolar relation in OW-FRMC framework considers multidimensional data on efficient identification of focal links with distance fuzzy classification measure, aiming at reducing the classification time. In addition, the bipolar relation results in efficient evaluation of acceptance and rejection measures and therefore improving the true positive rate. Weighted fuzzy measure of a multidimensional subset should be proportional to the length of that subset to sort the data streams with the higher classification ratio and reduce the computational complexity for classifying multi dimension data.

The OW-FRMC framework applies Correlative Significance Fuzzy Measure aiming at reducing the computational complexity for multi dimension data. With this, the optimal fuzzy rule mining classification framework maximizes the true positive rate and performs experimental evaluation on factors such as classification time, computational complexity for classifying multi dimension data and classification ratio. Experimental analysis shows that OW-FRMC framework is able to reduce the classification time for extracting the patterns by 13.55% and increase the true positive rate by 16.33% compared to the state-of-the-art works.

**Keywords:** Swift Rule, Piecewise Polynomial Modeling, Bipolar Relation, Weighted fuzzy measure, Multi dimension data.

## Introduction
Recent advancement in the pattern classification for efficient decision making process in business has opened the door towards fuzzy pattern classification. The traditional fuzzy rule mining classification classifies only the frequent events from the database. In the multidimensional data streams many research work is focused mainly on mining only the fuzzy rule but not focused on the fuzzy rule mining for multidimensional data.

Mining extensive classification rules for time series [1] using segmentation and piecewise polynomial modeling not only resulted in reducing energy consumption but also resulted in efficient classification. In [2] trajectories on road network were classified in an efficient manner using frequent pattern based classification improving the classification accuracy. Sentimental mining [3] using bit-map based approach on modern processors resulted in performance improvement. A generic local algorithm was presented in [4] that minimized the communication cost using decision tree and k-means clustering. Skyline and user's current navigational behavior in [5] with the application of effective ranking method extracted both static and dynamic preferences. However, the above said methods did not address classification ratio, which is addressed in OW-FRMC framework using weighted fuzzy rule. One of the major issues to be addressed in mining is security. In [6], Fast Distributed Mining (FDM) algorithm was applied using secure multi-party algorithm to reduce the computational cost and communicational round. In [7], random forest was applied for handling multiple target classes improving the computational complexity. Novel class detection using concept-drift [8] resulted in the improvement of stream classification accuracy. Application of suffix tree [9] for periodicity pattern mining that efficiently detected partial and full cycle time series. IR-Tree [10] for geographic document search handled both spatial and textual aspects resulted in the improvement of search time. However, the above methods did not address classification ratio that is provided in OW-FRMC framework using Weighted Fuzzy Rule.

Tracking movements is also receiving increasing attention with the introduction of mobile users and their applications. In [11], similarity of moving objects was determined based on the cluster ensembling algorithm and therefore reducing the energy consumption. Cluster based methods were introduced in [12] with the objective of efficient prediction of mobile users. Software Specification Discovery [13] shed light on software design using representative rules improving the efficiency. Subspace Clustering was applied in [14] for high dimensional data using region densities ensuring high quality. In [15] human behavior patterns were analyzed using Agglomerative Iterative Bayesian Fuzzy Clustering (AIBFC) for efficient prediction of human behavior. However, the computational complexity in the above said methods were not addressed, which is provided in OW-FRMC framework using weighted fuzzy measure.

Bayesian classifiers applied in [16] improved classification accuracy on large datasets with the aid of k-means clustering. In [17], fuzzy rough set approach was introduced to minimize the attributes that are redundant in nature using rule-based classifier. Sub graph pattern search [18] for graph streams resulted in efficient pruning using numerical vector space. In [19], concept-based user profiles were applied to obtain both positive and negative preference with the aim of improving the quality of resulting query. With the objective of improving classification accuracy, filter-based data partitioning was applied using feature-based and class-based measures.

Based on the aforementioned methods and techniques, in this paper, an effective framework called Optimal Weighted Fuzzy Rule Mining Classification (OW-FRMC) on filtered data stream is presented in the forthcoming sections.

## Optimal Weighted Fuzzy Rule Mining Classification

In this section, the framework Optimal Weighted Fuzzy Rule Mining Classification is introduced to classify multi dimension data called as OW-FRMC. The OW-FRMC framework designs a Weighted Fuzzy Rule Classification to improve the classification ratio by applying sequential rule to obtain the acceptance and rejection measure.

Then, with the introduction of Bipolar Relation for Multi dimension data, classification time is reduced due to the efficient classification of pattern based on multi dimension data. It also improves the true positive rate by applying threshold value to bipolar relation algorithm. Finally by applying Correlative Significance Fuzzy Measure, computational complexity involved in the classification of multidimensional data stream is reduced significantly.

## Design of Weighted Fuzzy Rule Classification

The first step in the design of Optimal Weighted Fuzzy Rule Mining Classification (OW-FRMC) on filtered data stream is the design of Weighted Fuzzy Rule Classification. The Weighted Fuzzy Rule Classification efficiently classifies the multidimensional data with bipolar property, aiming at improving the classification ratio. For efficient evaluation of Optimal Weight, two measures called confidence and support are used for efficient Fuzzy Rule Mining. Figure 1 given below shows the block diagram of Weighted Fuzzy Rule Classification.

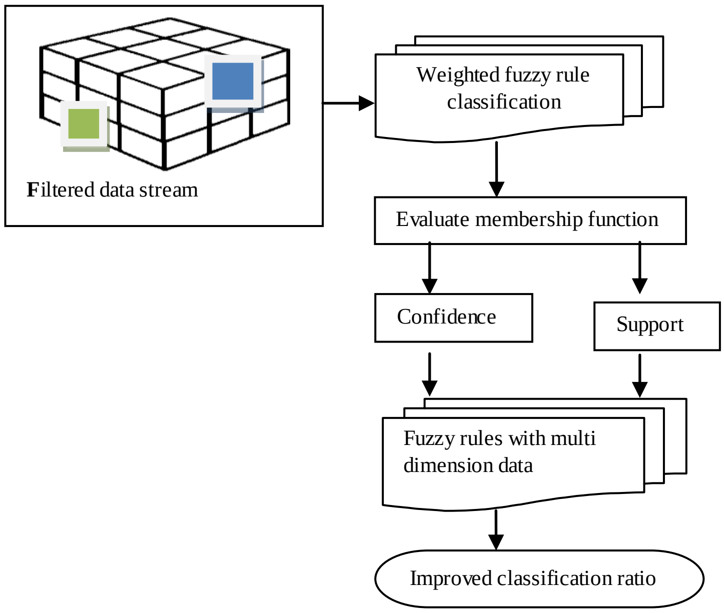


**Fig.1 Block diagram of Weighted Fuzzy Rule Classification**

As shown in figure 1, the block diagram of Weighted Fuzzy Rule Classification includes filtered data stream as input, where membership function is evaluated for it. Based on the membership function, the confidence and support value is measured. To the measured confidence and support values, fuzzy rules with multi dimension data are applied. This in turn improves the classification ratio.

Let us consider that there are '$n$' training patterns '$T = T_{p1}, T_{p2}, .., T_{pn}$' for an '$n\ dimension$' pattern classification problem. Then, the compatibility of training pattern '$T$' with respect to the antecedent '$A_p$' is formulated as given below.

$$\alpha_{A_p}(T) = \alpha_{A_{p1}}(T_{p1}), \ldots\ldots, \alpha_{A_{pn}}(T_{pn}) \in C \qquad (1)$$

From (1), '$\alpha_{A_p}(T)$' represents the membership function used in OW-FRMC on filtered data stream for antecedent fuzzy form '$T_{pi}$' in a class '$C$'. Based on the membership function, the confidence value is measured as formulated below.

$$Conf\left(A_p \rightarrow C_p\right) = \left(\frac{\sum_{T \in C_p} \alpha * A_p * T}{\sum_{T=1}^{n} \alpha * A_p * T}\right) \qquad (2)$$

Based on the membership function, the support value is derived as given below.

$$Supp\left(A_p \rightarrow C_p\right) = \left(\frac{\sum_{T \in C_p} \alpha * A_p * T}{n}\right) \qquad (3)$$

Using the confidence and support value, measured for each antecedent and consequent pattern, the proposed framework OW-FRMC using Optimal Weighted Fuzzy Rule Mining Classification, first explains fuzzy rules with a single dimension data. Let '$S$' represents a set of fuzzy rules. The OW-FRMC framework uses a sequential method for efficient classification of new patterns from the set of rules '$S$'. Then using sequential method, new rule '$Rule_{new}$' is evaluated for a new pattern and is formulated as given below.

$$\alpha\, A_{new}\ (T).\,Conf_{new} = MAX\,\{(\alpha * A_p T).\ (Conf_{A_p})\} \quad (4)$$

From (4), the new pattern '$T$' is classified as class '$Conf_{new}$' which forms the consequent class of the sequential rule. If several fuzzy rules have same maximum value but differing consequent classes for the new pattern '$T$', then the classification of '$T$' is rejected, otherwise the classification of '$T$' is accepted.

Now in the framework OW-FRMC using Optimal Weighted Fuzzy Rule Mining Classification, let us consider fuzzy rules with multiple dimension data. Then, the above equation is redefined as given below.

$$\alpha\, \mathbf{A}_{new}\ (\mathrm{T}).\,\mathrm{Conf}_{newy} = MAX\,\{(\alpha * A_p * T).\ (Conf_{A_{py}})\} \quad (5)$$

From (5), the new pattern '$T$' is classified as '$Conf_{newy}$' and the framework OW-FRMC reduces by defining '$Conf_{A_p}$' for each rule with multidimensional data as formulated below

$$Conf_{A_p} = MAX\left(Conf_{A_{py}}\right), where\ y = 1, 2, \dots, n \quad (6)$$

Thus from above (6), the rule weight of each fuzzy rule using single dimension data and multi dimension data shows can be efficiently used using Optimal Weighted Fuzzy Rule Mining Classification. This in turn improves the classification ratio. Figure 2 given below shows the algorithmic description of Optimal Weighted Fuzzy Rule (OWFR).

**Input:** Training patterns '$T = T_{p1}, T_{p2}, \dots, T_{pn}$', Antecedent '$A_p$'

**Output:** Improved classification rates on training patterns

Step 1: **Begin**
Step 2: **For** each training patterns '$T$'
Step 3: **For** each antecedent '$A_p$'
Step 4: Evaluate confidence using ()
Step 5: Evaluate support using ()
Step 6: Evaluate new pattern for single dimension data using ()
Step 7: Evaluate new pattern for multi dimension data using ()
Step 8: **End for**
Step 9: **End for**
Step 10: **End**

**Figure 2 OWFR algorithm**

The above OWFR algorithm is split into four steps. The first and second step measures the confidence and support value based on the membership function for different training patterns. The third step evaluates separately the new pattern for single dimension data based on the sequential rule. Finally, new patterns are classified for multi dimension data using sequential rule. This improves the classification ratio in a significant manner.

## Design of Bipolar Relation for Multi dimension data

The second step in the design of OW-FRMC framework is the design of bipolar relation for multi dimension data. The bipolar relation for multi dimension data in OW-FRMC framework measures acceptance and rejection measure. Using bipolar relation, to what extent a pattern should be included in the class and to what extent a pattern should be excluded in the class is considered. Using bipolar relation, the acceptance measure '$\beta_A^n\,(T)$' and rejection measure '$\beta_R^n\,(T)$' for each pattern '$T$'and a class '$C_i$' are evaluated.

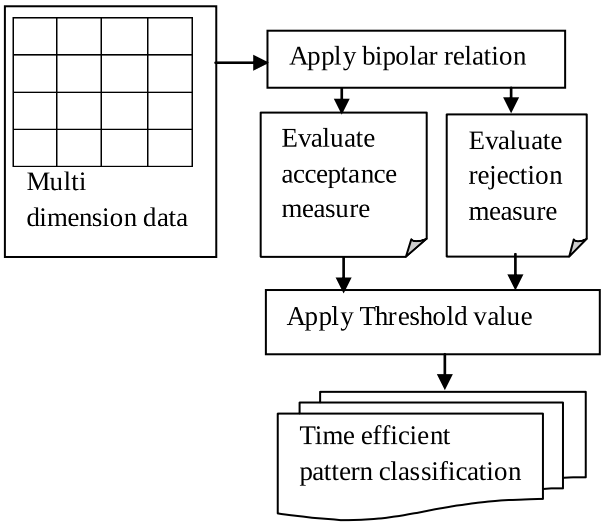


**Fig 3 Block diagram of Bipolar Relation for Multi dimension data**

As shown in figure 3, the bipolar relation in OW-FRMC framework is split into two zones called as inclusion zone (i.e. acceptance measure) and exclusion zone (i.e. rejection measure). Thus in order to characterize a pattern '$T$' characterized by vector '$V$' belongs to the class '$C_i$', the proposed OW-FRMC framework obtains the acceptance and rejection measure based on the feature '$f$'. Then, the acceptance and rejection measure is formulated as given below.

$$if\ fun_f^n\,(T) \in \beta_A^n\,(T),$$
$$then\ include\ pattern\ T\ in\ class\ 'C_i' \quad (7)$$

$if\ fun_f^n\ (T) \notin\ \beta_R^n\ (T)\ ,$
$then\ exclude\ pattern\ T\ from\ class\ C_i$' (8)

From (7) and (8), the factors leading to inclusion/exclusion of a pattern in/from class '$C_i$'' is efficiently formulated. Now, the extent to which a pattern should be included (i.e. acceptance measure) or excluded (i.e. rejection measure) is formulated using the bipolar relation with the aid of the threshold '$min$' and '$max$' based on the following conditions.

$$if\ \beta_A^n\ (T) \subseteq\ fun_{f,min}^n, then\ pattern\ is\ accepted \quad (9)$$

$$if\ \beta_A^n\ (T) \subseteq\ fun_{f,max}^n, then\ pattern\ is\ rejected \quad (10)$$

From (9) and (10), the acceptance measure and rejection measure using bipolar relation is efficiently evaluated. The bipolar relation in OWFRC function considers the multidimensional data on identifying the focal links with distance fuzzy classification measure.

The bipolar relation formulated the classification model that exists between the feature considered for pattern classification and the inclusion/exclusion of patterns in a specific class. As the humans make the efficient decision making process by considering the "pros" and "cons", the bipolar relation in OW-FRMC framework efficiently classifies the patterns based on the acceptance and rejection measure. As a result, the classification time for each pattern to be either in a class or not is done in an efficient manner. Figure shows the bipolar relation algorithm.

| |
|---|
| **Input:** Training patterns '$T = T_{p1}, T_{p2}, .., T_{pn}$',<br>Class '$Class_i = Cl_1, Cl_2, \dots, Cl_n$',<br>Threshold '$min, max$' |
| **Output:** Time efficient pattern classification |
| Step 1: **Begin**<br>Step 2: **For** each training patterns $T$<br>Step 3: **For** each class $Class_i$<br>Step 4: Perform inclusion function using ()<br>Step 5: Perform exclusion function using ()<br>Step 6: Based on the minimum threshold '$min$', pattern is accepted using ()<br>Step 7: Based on the maximum threshold '$max$', pattern is rejected using ()<br>Step 8: **End for**<br>Step 9: **End for**<br>Step 10: **End** |

**Figure 4 bipolar relation algorithm**

The algorithm for bipolar relation provided in figure performs two functions, inclusion and exclusion of patterns in a class. To start with based on the patterns and class, inclusion and exclusion function is performed. Then, based on the minimum and maximum threshold values '$\min and\ max$', either the pattern is accepted or rejected. This in turn minimizes the classification time for patterns by multiple attributes (i.e. multi dimension data) with respect to predefined classes in a significant manner.

## Correlative Significance Fuzzy Measure

Weighted fuzzy measure of multidimensional subset should be proportional to the length of the subset to sort the data streams with higher classification ratio. The OW-FRMC framework uses Correlative Significance Fuzzy Measure aiming at reducing the computational complexity while classifying multidimensional data. In the OW-FRMC framework, Correlative Significance Fuzzy Measure is introduced.

With the Correlative Significance Fuzzy Measure, the cardinality of multidimensional dataset is updated on the basis of the correlative significance, because certain patterns may be more or less important for the acceptance or rejection of a particular action. Let us consider that the correlative significance measure '$m_i$' with regards to pattern classification is obtained for each feature '$f$' from '$n$' elements, then the Correlative Significance Fuzzy Measure is measured as given below

$$\mu\ (T) =\ \left(\frac{T}{n}\right) *\ \sum_{i=1}^{N} w_i \qquad (11)$$

From (11), the Correlative Significance Fuzzy Measure '$\mu\ (T)$' is measured which further reduces the computational complexity during pattern classification. Based on results of equation (), given a pattern '$T$' characterized by its vector '$V$' and a class '$Cl_i$' with its features vector '$fun_f^n\ (T)$ (where $f = 1,2, \dots l$) and their correlative significance measure '$m_i$' the acceptance and rejection measure functions are determined as given below

$$\beta_A^n\ (T) =\ \gamma fun_A^n\ (T) * \left(\frac{f-1}{f}\right) *\ \ \mu\ (T) \quad (12)$$

$$\beta_R^n\ (T) =\ \gamma fun_R^n\ (T) * \left(\frac{f-1}{f}\right) *\ \ \mu\ (T) \quad (13)$$

From (12) and (13), the acceptance or rejection of a particular action for any pattern is obtained reducing the computational complexity for multidimensional data.

## Experimental setup

Optimal Weighted Fuzzy Rule Mining Classification (OW-FRMC) framework on multi dimensional data streams are performed in JAVA platform. The Java platform uses the Weka tool for the effective classification and effective classified the patterns from the filtered data streams. The OW-FRMC framework uses the Japanese Vowels multi dimensional Data Set from UCI repository to perform the experimental work. Japanese Vowels multi dimensional dataset records 640 time series of 12 Linear Predictive coding (LPC) from nine male speakers.

The collected multidimensional data are used for classification of efficient patterns on the basis of irrelevant Japanese vowels attributes from data streams. For each occurence, the analysis constraint explains 12-degree linear prediction analysis to get hold of discrete-time series with 12 LPC cepstrum coefficients. Each utterance by a speaker forms a time series in the range of 7-29 and each position of a time series is of 12 features (i.e.,) coefficients, with the total time series about 640.

The OW-FRMC framework is compared against the existing Mining Comprehensible Classification Rules for Time Series (MCCR-TS) [1] and Mining Patterns for Classifying Trajectories (MP-CT) [2]. The experiment is conducted on the factors such as classification ratio, classification time, true positive rate and computational complexity for classifying multidimensional data.

**Discussion**

The performance of Optimal Weighted Fuzzy Rule Mining Classification (OW-FRMC) framework is compared with the existing Mining Comprehensible Classification Rules for Time Series (MCCR-TS) [1] and Mining Patterns for Classifying Trajectories (MP-CT) [2]. The performance is evaluated according to the following metrics.

## Impact of classification ratio

The classification ratio is the ratio of difference between the total items and the items not properly classified to the total items. The mathematical formulation for classification ratio is given as below.

$$CR = \left(\frac{n - Items\ not\ properly\ classified}{n}\right) * 100 \qquad (14)$$

From (14), the classification ratio '$CR$' is measured based on the total number of items '$n$' and is measured in terms of percentage (%). Higher the classification ratio, more efficient the method is said to be. The table 1 represents the classification ratio obtained using JAVA platform and comparison is made with two other methods, namely MCCR-TS [1] and MP-CT [2]. To conduct experiments, 175 items were considered.

## Table 1 Tabulation for classification ratio

| Number of items (n) | Classification ratio (%) | | |
|---|---|---|---|
| | OW-FRMC | MCCR-TS | MP-CT |
| 25 | 73.15 | 61.48 | 41.89 |
| 50 | 75.19 | 70.16 | 64.11 |
| 75 | 79.24 | 74.21 | 68.16 |
| 100 | 71.35 | 66.32 | 60.27 |
| 125 | 74.26 | 69.23 | 63.18 |
| 150 | 78.55 | 73.52 | 67.47 |
| 175 | 82.13 | 75.10 | 69.05 |

Figure 5 shows the result of classification ratio versus the varying number of items. To better perceive the efficiency of the proposed OW-FRMC framework, substantial experimental results are illustrated in Figure 5. The OW-FRMC framework is compared against the existing MCCR-TS [1] and MP-CT [2].

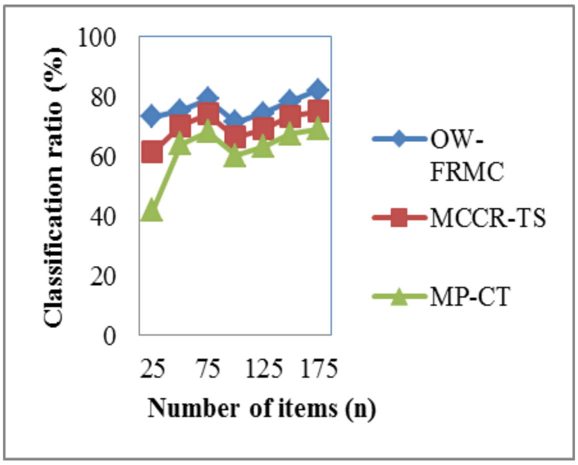


**Figure 5 Measure of classification ratio**

Results are presented for different number of items. Higher, the number of items, more successful the method is. The results reported here confirm that with the increase in the number of items, the classification ratio also increases, though betterment achieved using OW-FRMS framework. The process is repeated for 175 different items for conducting experiments.

As illustrated in Figure 5, the proposed OW-FRMS framework performs relatively well when compared to two other methods MCCR-TS [1] and MP-CT [2]. The OW-FRMS framework had better changes using the extensive Weighted Fuzzy Rule Classification. This is because in order to obtain better classification ratio, a bipolar property is applied to the multidimensional data that results in the improvement of classification ratio. Furthermore, based on the membership function, the support and confidence is evaluated using optimal weight which in turn improves the classification ratio by 8.25% compared to MCCR-TS and 18.84% compared to MP-CT.

**Impact of classification time**

Classification time is the time taken to classify the patterns with multi dimensional data. Classification time is the product of time taken to classify single pattern with respect to total number of patterns taken for experimental purpose.

$$CT = T * Time\ (single\ pattern) \qquad (15)$$

From (15), the classification time '$CT$' is measured with respect to the number of patterns '$T$' and is measured in terms of milliseconds (ms). Lower the classification time, more efficient the method is said to be.

**Table 2 Tabulation for classification time**

| Number of patterns (T) | Classification time (ms) | | |
|---|---|---|---|
| | OW-FRMC | MCCR-TS | MP-CT |
| 5 | 12.84 | 17.25 | 20.13 |
| 10 | 18.35 | 26.37 | 31.40 |
| 15 | 28.31 | 36.33 | 41.36 |
| 20 | 41.25 | 49.27 | 54.30 |
| 25 | 35.86 | 43.88 | 48.91 |
| 30 | 43.16 | 51.18 | 56.21 |
| 35 | 55.29 | 63.31 | 68.34 |

In order to decrease the classification time with respect to different number of patterns, the classification time using the framework OW-FRMC and two methods, MCCR-TS and MP-CT are presented with visual comparison in table 2. The results for 35 different patterns are illustrated in figure 6. The classification time to classify the patterns using our framework OW-FRMC offer comparable values than the state-of-the-art methods.

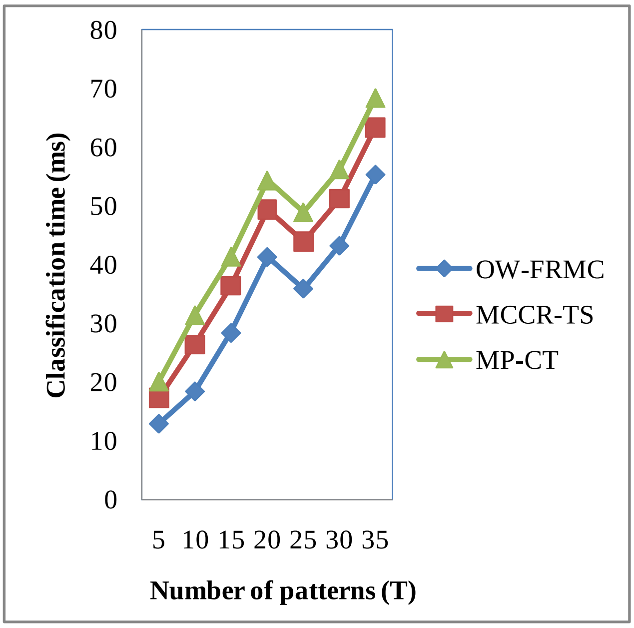


**Figure 6 Measure of classification time**

The targeting results of classification time using OW-FRMC framework is compared with two state-of-the-art methods [1], [2] in figure 6 is presented for visual comparison based on the number of patterns. Our framework differs from the MCCR-TS [1] and MP-CT [2] in that we have incorporated bipolar relation for multidimensional data. By applying bipolar relation for multidimensional data, the acceptance and rejection measure are evaluated by applying a threshold value. In addition, the extent to which a pattern should be included in the class and the extent to which a pattern should be excluded in the class are considered with the aid of the threshold '$min$' and '$max$'. This threshold value in OW-FRMC framework efficiently classifies the focal links with distance fuzzy classification measure. Therefore the classification time is reduced by 25.89% compared to MCCR-TS and 42.26% compared to MP-CT respectively.

## Impact of true positive rate

In a classification task on multidimensional data stream, the true positive rate for a class is the number of patterns correctly classified as true positives divided by the total number of elements. The mathematical formulation for true positive rate is as given below.

$$TPR = \left(\frac{Number\ of\ patterns\ correctly\ classified}{T}\right) * 100 \qquad (16)$$

From (16), the true positive rate '$TPR$' is obtained with respect to the patterns '$T$' and measured in terms of percentage (%).The comparison of true positive rate time for pattern classification is presented in table 3 with respect to different patterns in the range of 5 to 35. With increase in the number of patterns, the true positive rate for pattern classification also gets increased.

**Table 3 Tabulation for true positive rate**

| Number of patterns (T) | True positive rate (%) | | |
|---|---|---|---|
| | OW-FRMC | MCCR-TS | MP-CT |
| 5 | 80.35 | 62.45 | 43.82 |
| 10 | 82.19 | 74.16 | 69.11 |
| 15 | 75.28 | 67.25 | 58.20 |
| 20 | 84.56 | 76.53 | 71.48 |
| 25 | 86.31 | 78.28 | 73.23 |
| 30 | 78.29 | 70.26 | 65.21 |
| 35 | 90.35 | 82.32 | 75.27 |

Figure 7 given below shows the true positive rate for efficient pattern classification for OW-FRMC framework, MCCR-TS [1] and MP-CT [2] versus increasing number of patterns in the range of 5 to 35. The true positive rate improvement returned over MCCR-TS and MP-CT increase gradually as the number of patterns gets increased though not linear because of the changes observed in the patterns of measure.

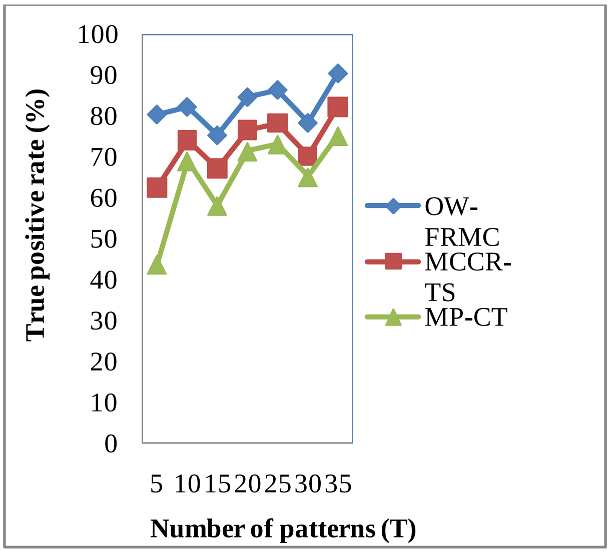


**Figure 7 Measure of true positive rate**

From figure 7, it is illustrative that the true positive rate for pattern classification is reduced using the proposed framework OW-FRMC. For example when the number of patterns is 20, the percentage improvement of OW-FRMC framework compared to MCC-TS is 9.49 percent and compared to MP-CT is 15.46 percent respectively. This is because the bipolar relation by using the focal links with distance fuzzy classification measure efficiently evaluates the patterns that should be accepted and those patterns that should be rejected. Therefore, the true positive rate is improved by 11.52% compared to MCC-TS. Furthermore, by applying bipolar relation algorithm, OW-FRMC framework obtains the acceptance and rejection measure based on the feature '$f$' and therefore improving the true positive rate by 21.15% compared to MP-CT respectively.

## Impact of computational complexity

Computational complexity is the time taken to evaluate the computational process with regards to multidimensional data. Computational complexity is the product of time taken to compute single item with respect to total number of items considered for experimental purpose.

$$CC = n * Time\ (single\ item) \tag{17}$$

From (17), the computational complexity '$CC$' is evaluated with respect to the number of items '$n$' and time for single item respectively. Lower the computational complexity, more efficient the method is said to be and is measured in terms of milliseconds (ms).

**Table 4Tabulation for computational complexity**

| Number of items (n) | Computational complexity (ms) | | |
|---|---|---|---|
| | OW-FRMC | MCCR-TS | MP-CT |
| 25 | 33.15 | 40.26 | 47.31 |
| 50 | 41.28 | 47.31 | 52.36 |
| 75 | 55.87 | 61.90 | 66.95 |
| 100 | 61.32 | 67.35 | 72.40 |
| 125 | 59.45 | 66.48 | 71.53 |
| 150 | 64.13 | 70.16 | 75.21 |
| 175 | 71.35 | 77.38 | 82.43 |

The computational complexity for computational processing for OW-FRMC framework is elaborated in table 4 and comparison made with two other methods MCCR-TS and MP-CT respectively. We consider the framework with 175 items for experimental purpose using JAVA.

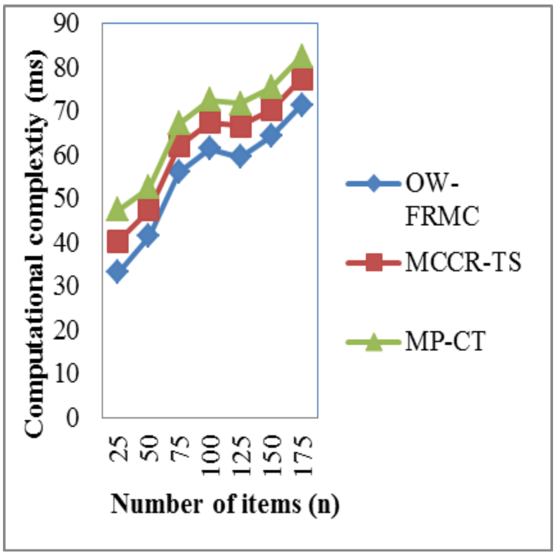


**Figure 8 Measure of computational complexity**

Figure 8 shows the measure of computational complexity with respect to differing number of items. The computational complexity using OW-FRMC framework is reduced owing to the fact that the proposed framework uses Correlative Significance Fuzzy Measure. With this Correlative Significance Fuzzy Measure, the acceptance and rejection measure functions are determined for any pattern reducing the computational complexity by 12.33% compared to MCCR-TS and 22.94% compared to MP-CT respectively.

.

## Conclusion

In this work, an effective framework called Optimal Weighted Fuzzy Rule Mining Classification (OW-FRMC) is presented. The framework reduces the computational complexity with respect to multidimensional data with minimum classification time for pattern classification and therefore provides higher classification accuracy. The goal of our optimal weighted fuzzy rule mining classification is to improve the classification ratio using the training and test patterns obtained from the Japanese Vowels multi dimensional Data Set from UCI repository which significantly contribute to the relevance. To do this, we first designed a Weighted Fuzzy Rule Classification to determine the confidence and support value based on the membership function to improve the classification ratio. Then, based on this measure, we proposed a Bipolar Relation for Multi dimension data that efficiently obtains the acceptance and rejection measure reducing the classification time and improving the true positive rate in an extensive manner. With the acceptance and rejection measure, the cardinality of multidimensional dataset is updated on the basis of the correlative significance reducing the computational complexity for pattern classification. In addition Optimal Weighted Fuzzy Rule algorithm and bipolar relation algorithm ensures efficient pattern classification for varied training and test patterns. Through the experiments using Japanese Vowels multi dimensional Data Set from UCI repository, we observed that the pattern classification provided more accurate results compared to existing methods. The results show that OW-FRMC framework offers better performance with an improvement of classification ratio by 13.55% and reduces the computational complexity by 17.63% compared to MCCR-TS and MP-CT respectively.

## REFERENCES

[1] Dominik Fisch, Thiemo Gruber and Bernhard Sick, “SwiftRule: Mining Comprehensible Classification Rules for Time Series Analysis”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 5, May 2011, Pages 774 – 787.

[2] Jae-Gil Lee, Jiawei Han, Xiaolei Li and Hong Cheng, “Mining Discriminative Patterns for Classifying Trajectories on Road Networks”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 5, May 2011, Pages 713 – 726.

[3] Morten Middelfart, Torben Bach Pedersen, and Jan Krogsgaard, “Efficient Sentinel Mining Using Bitmaps on Modern Processors”, IEEE Transactions on Knowledge and Data Engineering, Volume 25, Issue 10, October 2013, Pages 2231 – 2244.

[4] Ran Wolff, Kanishka Bhaduri ,and Hillol Kargupta, “A Generic Local Algorithm for Mining Data Streams in Large Distributed Systems”, IEEE Transactions on Knowledge and Data Engineering, Volume 21, Issue 4, April 2009, Pages 465 – 478.

[5] Zhiyuan Chen, Tao Li, and Yanan Sun, “A Learning Approach to SQL Query Results Ranking Using Skyline and Users’ Current Navigational Behavior”, IEEE Transactions on Knowledge and Data Engineering, Volume 25, Issue 12, December 2013, Pages 2683 – 2693.

[6] Tamir Tassa, “Secure Mining of Association Rules in Horizontally Distributed Databases”, IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 4, April 2014, Pages 970 – 983.

[7] Hanady Abdulsalam, David B. Skillicorn, and Patrick Martin, “Classification Using Streaming Random Forests”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 1, January 2011, Pages 22 – 36.

[8] Mohammad M. Masud, Jing Gao, Latifur Khan, Jiawei Han, and Bhavani Thuraisingham, “Classification and Novel Class Detection in Concept-Drifting Data Streams under Time Constraints”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 6, June 2011, Pages 859 – 874.

[9] Faraz Rasheed, Mohammed Alshalalfa, and Reda Alhajj, “Efficient Periodicity Mining in Time Series Databases Using Suffix Trees”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 1, January 2011, Pages 79 – 94.

[10] Zhisheng Li, Ken C.K. Lee, Baihua Zheng, Wang-Chien Lee, Dik Lun Lee, and Xufa Wang, “IR-Tree: An Efficient Index for Geographic Document Search”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 4, April 2011, Pages 585 – 599.

[11] Hsiao-Ping Tsai, De-Nian Yang, and Ming-Syan Chen, “Mining Group Movement Patterns for Tracking Moving Objects Efficiently”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 2, February 2011, Pages 266 – 281.

[12] Eric Hsueh-Chan Lu, Vincent S. Tseng, and Philip S. Yu, “Mining Cluster-Based Temporal Mobile Sequential Patterns in Location-Based Service Environments”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 6, June 2011, Pages 914 – 927.

[13] David Lo, Jinyan Li, Limsoon Wong, and Siau-Cheng Khoo, “Mining Iterative Generators and Representative Rules for Software Specification Discovery”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 2, February 2011, Pages 282 – 296.

[14] Yi-Hong Chu, Jen-Wei Huang, Kun-Ta Chuang, De-Nian Yang, “Density Conscious Subspace Clustering for High-Dimensional Data”, IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 1, January 2010, Pages 16 – 30.

[15] Sang Wan Lee, Yong Soo Kim, and Zeungnam Bien, “A Nonsupervised Learning Framework of Human Behavior Patterns Based on Sequential Actions”, IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 4, April 2010, Pages 479 – 492.

[16] Carlos Ordonez and Sasi K. Pitchaimalai, “Bayesian Classifiers Programmed in SQL”, IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 1, January 2010, Pages 139 – 144.

[17] Suyun Zhao, Eric C.C. Tsang, Degang Chen, and XiZhao Wang, “Building a Rule-Based Classifier—A Fuzzy-

Rough Set Approach", IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 5, May 2010, Pages 624 – 638.

[18] Lei Chen, and Changliang Wang, "Continuous Sub graph Pattern Search over Certain and Uncertain Graph Streams", IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 8, August 2010, Pages 1093 – 1109.

[19] Kenneth Wai-Ting Leung and Dik Lun Lee, "Deriving Concept-Based User Profiles from Search Engine Logs", IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 7, July 2010, Pages 969 – 982.

[20] Rozita A. Dara, Masoud Makrehchi, and Mohamed S. Kamel, "Filter-Based Data Partitioning for Training Multiple Classifier Systems", IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 4, April 2010, Pages 508 – 522.