# Fuzzy Association Rule Mining for Wireless Sensor Network Data

**Manisha Rajpoot and H R Sharma**

Department of Computer Science and Engineering,
Rungta College of Engineering and Technology, Bhilai, India

**Abstract:** Association rule mining in a wireless sensor network data is an important research area. This data are given the opportunity to deal with the uncertainty of data. Fuzzy association rule mining can be an important method in wireless sensor network data due to its the completeness and able to handle the uncertainty. In this work, a fuzzy association rule mining is used for wireless sensor networks' data. Our experiment shows that fuzzy association rule mining in wireless sensor networks' data gives better performance as non-fuzzy association rule mining. It performs in less time and generates less frequent set compare as existing algorithm.

**Keywords:** Wireless Sensor Network, Association Rule Mining, Fuzzy Set

## 1. Introduction

Wireless sensor networks (WSNs) are becoming ubiquitous; a set of broad requirements is beginning to emerge across high-priority applications including disaster preparedness and management, adaptability to climate change, national or homeland security, and the management of critical infrastructures. The raw data from WSNs need to be efficiently managed and transformed into usable information through data mining. Association rule mining for WSNs have received lots of attention due to their importance. An example of such a rule is (s1s2 ⇒ s3, 75%, λ), which can be interpreted as follows: If the events from sensor s1 and s2 are received, then there is an 80% chance to receive an event from sensor s3 within λ unit of time where 75% is the frequency of the rule [1][2][13].

WSNs data in real-world applications, however, usually consist of fuzzy and quantitative values. Designing sophisticated data-mining algorithms able to deal with various types of data presents a challenge to workers in this research field. Fuzzy set theory was first proposed by Zadeh [3]. It has been used more and more frequently in intelligent systems because of its simplicity and similarity to human reasoning. It is primarily concerned with quantifying and reasoning using natural language in which words can have ambiguous meanings. This can be thought of as an extension of traditional crisp sets, in which each element must either be in or not in a set. The theory has been applied in many fields such as manufacturing, engineering, diagnosis, economics, among others. Several fuzzy learning algorithms for inducing rules from giving sets of data have been designed and used with good effects in specific domains. The fuzzy set theory has been applied to find interesting association rules or sequential patterns in transaction data with quantitative values [4]. In this work a fuzzy association will be applied to WSNs data.

The remainder of the paper is organized as follows: In Section 2, related work in pattern mining for WSNs data will be reported. Section 3 Framework for Fuzzy Association Rule in WSNs will be given. The experiment and results analysis will be reported in Section 4 and finally, study will be concluded in Section 5.

## 2. Related Works

In this section, we review related works that have been proposed for mining association rules to sensor data. Recently, several works have been proposed for applying association rules on the WSN data [1][2][5][6][7]. Most of these works have targeted the data values of the sensor nodes; in other words, the values provided by the sensors are the main objects of the rules [14]. In these studies, the time is divided into intervals, and the sensor's values at that interval formulate the context to be stored in the database. Each different value of a sensor is regarded as a single element, and it is assumed that sensors take on a finite number of discrete states.

Romer [8] proposed an in-network data mining technique to discover frequent patterns of events with certain spatial and temporal properties.

In [6], the authors have proposed sensor association rules in which the event-detecting sensors are the main objects of the rules regardless of their values. To store the sensor's event- detecting status, this method uses a representation, called a positional lexicographic tree (PLT), constructed in lexicographic order of sensors. Each sensor is mapped to a unique integer, so that the lexicographic order is preserved. The first step in constructing the PLT is to scan the database once to obtain the set of frequent event–detecting sensors. The set of sensors that detect event at the same unit of time is processed

(DOI: dx.doi.org/14.9831/0971-9563.2014/2-4/BRIS.16)

together as an entry in the PLT. With the second database scan, the frequent sensors on each of such set are transformed into a position vector constructed by mapping the lexicographic distance between a sensor's identifier, and its parent's identifier, and then the vector is inserted into the PLT. This approach constructs as many PLTs as the number of distinct last sensors of all position vectors. The construction of PLTs terminates when all the position vectors from the database are inserted into respective PLT structures. Similar to the FP-growth approach, PLT follows a pattern growth mining technique.

The mining begins Tanbeer SK, et al. [9] Single-pass Algorithm for Mining Association Rules from Wireless Sensor Networks with the sensor having the maximum rank by generating the frequent patterns from its PLT in a recursive way. A conditional vector considering only the prefix part in the PLT for the sensor (under mining) is constructed. At this stage, the PLT of the conditional vector, if available, is also updated. For all of the sensors present in PLT structures, the mining process is the same. The computation required at each recursion to update the PLT involved in the prefix part of a pattern is not trivial. Therefore, the two database scans requirement and the additional PLT update operations during mining limit the efficient use of this approach in handling WSN data.

Salah et al. [10] proposed enhancement of T-pattern approach to mine for temporal patterns in sensor data. The compression based techniques and independence testing were used to improve the performance.

Samarah et al.[5] Proposed target association rules for point of coverage, wireless sensor networks. This algorithm is used to discover correlation among a set of target monitored by a WSNs. AI-node, Schedule-buffer and Fused schedule buffer mechanism were used in this algorithm.

## 3. Framework for fuzzy association rule in WSNs

Let S = {$s_1$, $s_2$, …, $s_n$) be a set of sensors, in a particular sensor network. It is assumed that the time space is divided into equally sized slots {$t_1$, $t_2$,.., $t_m$} such that $t_{i+1} - t_i = \lambda$, $i \in [1, m- 1]$ and $\lambda$ is the size of each time slot.

A set P = {$s_1$, $s_2$… $s_k$} $\subseteq$ S is called a pattern of sensors. A sensor database, SD, is defined to be a set of epochs in which each epoch is coupled E ($E_{ts}$, A) such that A is a pattern of event detecting sensors that report events within the same time slot. $E_{ts}$ is the epoch's time slot. Let size (E) be the size of E, i.e., the number of sensors in A. If each quantitative attribute $A_j$ is extended by its fuzzy set, we can get the extended attribute set $I_f$ from P. Using the corresponding membership functions defined with each fuzzy set, the original data set SD is changed into a fuzzy dataset $SD_f$.

We say an epoch $E_f$ ($E_{tsf}$, $A_f$) supports a pattern $A_f$' if $B_f \subseteq A$. The frequency of the pattern $B_f$ is $SD_f$ is defined to be the number of epochs in $SD_f$. The standard approach to evaluate the significance of fuzzy association rules in WSNs is to extend the definition of support and confidence measures to fuzzy association rules [11]. The degree of support of the rule $A_f \Rightarrow B_f$ for the whole $SD_f$ is defined as:

$$SD\ conf\ \left(A_f \Rightarrow B_f\right) = \frac{\sum_{i=1}^{n} A_f(x) \otimes B_f(y)}{|SD_f|}$$

And the degree of confidence is defined as:

$$SD\ conf\ \left(A_f \Rightarrow B_f\right) = \frac{\sum_{i=1}^{n} A_f(x) \otimes B_f(y)}{A_f(x)}$$

where $|SD_f|$ is the total number of transactions in $SD_f$ , which is equal to N, the number of transactions in the quantitative database SD. $A_f(x)$ and $B_f(y)$ denote the degree of membership of the elements x and y with respect to the fuzzy set $A_f$ and $B_f$, respectively, $\otimes$ is a t-norm [8]. Based upon the notations of SDsupp and SDconf, a rule $A_f \Rightarrow B_f$ is the interesting fuzzy association rule if

1. SDsupp($A \Rightarrow B$)$\geq$ min_supp;
2. SDconf($A \Rightarrow B$) $\geq$ min_conf;

Where min_supp and min_conf are the thresholds defined by users.

## 4. Experiment and Results

Both the association rule mining algorithms fuzzy and non- fuzzy for WSNs were implemented in a Java program language. It experimented on a 3.20 GHz Pentium processor, 4 GB RAM and window 2007 operating system. The experiments were conducted with the real data sets (chess and connect-4) [12]. These datasets also earlier used to compare the performance of pattern mining in wireless sensor network by other authors [7]. Fuzzy ARM experimented with these data sets and Table 1 shows the execution time and generated the number of frequent set and association rule from the fuzzy ARM. The experiments were conducted with constant confidence value 80% with variable minimum support values. Figure1 and 2 shows a comparison between WSNs ARM and WSNs Fuzzy ARM basis of execution time and number of frequent sets generated by these algorithms. The execution time and number of frequent set was taken in log10 form to accommodate values within the graph in a proper way. The experiments show that the Fuzzy ARM produces better results. It can be observed from

(DOI: dx.doi.org/14.9831/0971-9563.2014/2-4/BRIS.16)

figure 1 that WSNs fuzzy ARM is taking less time compare to WSNs ARM. Further, WSNs fuzzy ARM generated less frequent sets compare to WSNs ARM.

**Table 1**. Experiment Result of WSNs fuzzy ARM

| Support % | Time (s) | | No. of Frequent set | | No. of rules | |
|---|---|---|---|---|---|---|
| | *Chess* | *Connect* | *Chess* | *Connect* | *Chess* | *Connect* |
| 80 | 0.06 | 2.13 | 726 | 2438 | 3678 | 13184 |
| 70 | 0.08 | 2.53 | 1499 | 3784 | 65584 | 18702 |
| 60 | 0.11 | 2.97 | 2748 | 5312 | 10108 | 23048 |
| 50 | 0.13 | 3.32 | 4568 | 6956 | 14413 | 27027 |
| 40 | 0.16 | 3.64 | 6824 | 8824 | 18061 | 31023 |

**Table 2**. Experiment Result of WSNS Fuzzy ARM And Non Fuzzy ARM on data set Chess

| Support % | Time (s) | | No. of Frequent set | |
|---|---|---|---|---|
| | *Fuzzy ARM* | *ARM* | *Fuzzy ARM* | *ARM* |
| 80 | 0.06 | 3.33 | 726 | 8227 |
| 75 | 0.09 | 6.22 | 1067 | 20993 |
| 70 | 0.09 | 13.59 | 1499 | 48731 |
| 65 | 0.11 | 39.69 | 2059 | 111239 |
| 60 | 0.11 | 75.57 | 2748 | 254944 |
| 55 | 0.12 | 175.19 | 3603 | 574998 |
| 50 | 0.13 | 350.14 | 4568 | 1272932 |

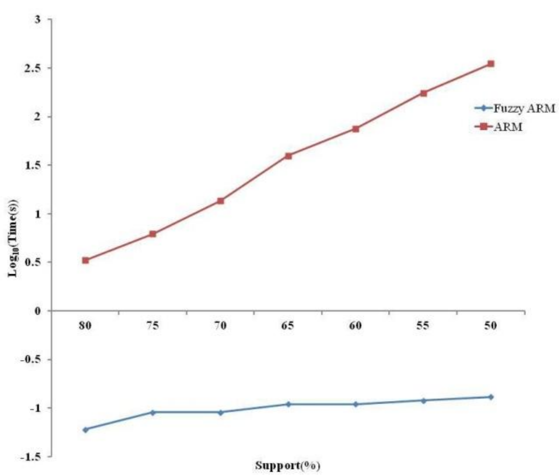


**Fig. 1.** Comparison based on execution time of Fuzzy ARM and ARM with Chess data set.

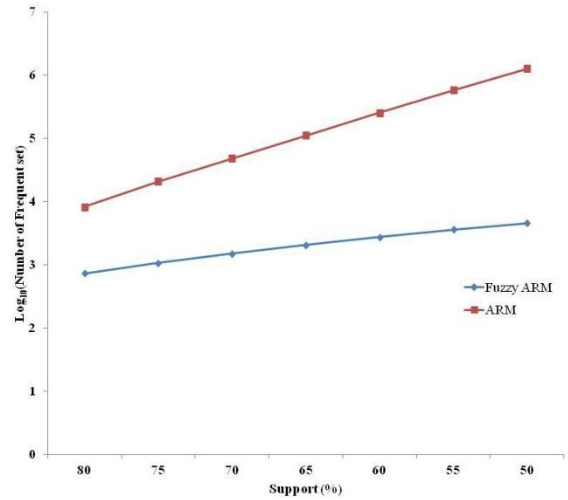


**Fig. 2.** Comparison based on the number of frequent sets of Fuzzy ARM and ARM with chess data set

## 5. Conclusion

In this paper, we presented a framework of fuzzy ARM algorithm to mine the Wireless Sensor Networks' data. The fuzzy ARM was implemented and test with data sets. The experiment illustrated that fuzzy ARM is more efficient as traditional ARM for WSNs. Further, it incorporated the features of fuzzy set such handle the uncertainty of data. It is assumed from the experimental result that this fuzzy technique will be an alternative technique to generate patterns and association rule from WSNs data.

## References


1. Farzanyar Z and Kangavari M. Efficient mining of fuzzy Association rules from The Pre-processed data set. Computing and Informatics. 2012; 31: 331–347
2. Boukerche A and Samarah S. An Efficient Data Extraction Mechanism for Mining Association Rules from Wireless Sensor Networks. Proc. IEEE International Conference on Communications, 2007: 3936-941.
3. Zadeh LA. Fuzzy sets. Information and Control. 1965; 8(3): 338–353.
4. Hong TP and Lee YC. An Overview of Mining Fuzzy Association Rules" H. Bustince et al. (Eds.), Fuzzy sets and Their Extension: Representation, Aggregation and Models. 2008; 220: 397-410.
5. Samarah S, Boukerche A and Habyalimana AS. Target Association Rules: A New Behavioral Patterns for Point of Coverage Wireless Sensor Networks. IEEE Transactions on Computers. 2011; 60(6): 879-889.
6. Boukerche A and Samarah S. A Novel Algorithm for Mining Association Rules in


(DOI: dx.doi.org/14.9831/0971-9563.2014/2-4/BRIS.16)

Wireless Ad Hoc Sensor Networks. IEEE Transactions on Parallel and Distributed Systems. 2008; 19(7): 865-877.
7. Tanbeer SK Ahmed CF and Jeong BS. An Efficient Single-Pass Algorithm for Mining Association Rules from Wireless Sensor Networks. IETE Technical Review. 2009; 26(4):280-289.
8. Romer K. Distributed Mining of Spatio-Temporal Event Patterns in Sensor Networks. Proc. International Conference on Distributed Computing in Sensor Systems / Euro-American Workshop Middleware for Sensor Networks (DCOSS/ EAWMS), 2006; 103-16.
9. Tanbeer SK, Ahmed CF, Jeong BS and Lee YK. Efficient single-pass frequent pattern mining using a prefix-tree. Information Sciences. 2009; 179(5): 559-83.
10. Salah AA et al. T-Patterns Revisited: Mining for Temporal Patterns in Sensor Data, Sensors 2010; 10: 7496-7513.
11. Sheibani R and Ebrahimzadeh A. An Algorithm For Mining Fuzzy Association Rules. Proc. of the International MultiConference of Engineers and Computer Scientists. 2008.
12. Blake CL and Merz CJ. UCI repository of machine learning databases, University of California – Irvine, Irvine, CA, 1998.
13. Rashid M, Gondal I and Kamruzzaman J. Mining associated sensor patterns for data stream of wireless sensor networks, PM2HW2N '13 Proceedings of the 8th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks, 2013; 91-98.
14. Mahmood A et al. Data Mining Techniques for Wireless Sensor Networks: A Survey. International Journal of Distributed Sensor Networks, 2013; Article ID 406316, 24 pages.

(DOI: dx.doi.org/14.9831/0971-9563.2014/2-4/BRIS.16)