Volume 5, Issue 11, November 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

ASPMS Based Approach to Mine Frequent Pattern from WSN Dataset

Snehal Rewatkar, Amit Pimpalkar
Department of CSE, G. H. Raisoni Academy of Engineering & Technology
Nagpur, Maharashtra, India

Abstract— ***Sensors produce huge amounts of data during transaction in network and those data transfer to the base station that consume more energy because long time of transmission. In this paper, we propose Associated Sensor Pattern Mining of data Stream (ASPMS) algorithm, which can find frequent patterns among sensors in Wireless Sensor Networks (WSNs) and also improving the WSNs quality of service. This algorithm can extracts knowledge from sensors data. ASPMS identify sets of correlated temporally sensors with uses an innovative tree structure called associated sensor pattern stream tree (ASPS tree). ASPS-tree and algorithm is used to generate associated patterns with one scan database. ASPMS will perform using sliding window-based associated sensor pattern mining for WSNs. It can utilize less memory as it compresses the same frequency nodes into a single node using Branch Sort Method (BSM). The experimental result of the proposed algorithm will give better result than existing algorithm.***

Keywords— WSN, ASPMS, ASPS-tree, BSM.

I. INTRODUCTION

WSNs are being used for diverse applications such as low cost area monitoring, structural monitoring and military surveillance, etc. Data mining techniques have been used to extract useful knowledge from WSN data, through discovering relationships among the sensor nodes which are known as behavioural patterns. More recently, research has been focused to mine different types of behavioural patterns, e.g., sensor association rules from stored (static) sensor data context association rules from sensor data stream, associated sensor patterns and regularly frequent sensor patterns from static as well as stream data.

The existing system Share-Frequent Sensor Pattern (SFSP) find interesting knowledge from sensor databases or streams by considering the non-binary frequency. The discovery of SFSP from WSNs can be useful in a variety of real-world applications that require the fine-grain monitoring of physical environments such as bridges, fire, toxic gas leaks and explosion. SFSP capture the temporal relationships among the sensor nodes during their event detection process. Therefore, on occurrence of an event associated sensors as revealed by mining results can be activated or certain. From that work they have to find frequent patterns by analysing wide range of used algorithms with the purpose of to discover and to obtain frequent patterns over large databases.

Mining of sensor data for extraction the useful knowledge is a difficult task. In existing works generate sensor association rules to extract the knowledge for frequency patterns. In this paper, propose a new type of pattern called ASPMS which capture correlated association as well as temporal correlations. To extract such patterns, it uses a tree structure called ASP-tree that is based upon mining algorithm and stores sensor data in manner. ASPS-tree and algorithm is used to generate associated patterns with one scan database as compare to Apriori and Frequent Pattern(FP) Growth algorithm. The efficiency result of algorithm has been shown in comparative study: Apriori algorithm, FP Growth algorithm and ASPMS frequent pattern mining algorithms.

II. RELATED WORK

Several mining techniques have been proposed in literature to mine transaction data from sensors in network. M. Rashid, et.al [2] introduced behavioural pattern called share frequent sensor patterns (SFSPs) extracted the temporal relations and more useful knowledge from WSNs data. Share-frequent sensor pattern tree (ShrFSP-tree) used to discover such pattern that avoids level-wise candidate generation. It also present a parallel and distributed framework that reduced the I/O cost with inter-processor communication. This method shows efficiency of time and memory to find SFSPs than existing.

Fast Share Measure (FSM) algorithm studied in [3] generated all frequent itemsets. FSM work with level closure property that decrease itemsets to be counted .The results of FSM algorithm is superior to the zero subset pruning (ZSP) algorithm between magnitude of 0.2% and 2% minimum thresholds. In [5], introduced a parallel and distributed mining algorithm to solve the problem of parallel and distributed computing system based on FP-tree structure, Load Balancing FP-Tree (LFP-tree). LFP evaluate the function of loading each item. The results show, LFP-tree has better speed-up and can reduce the computation time compared with Parallel FP-Tree (PFP-tree).

Incremental Share-Frequent Pattern Tree (IncrShrFP-Tree) [9] avoided the level wise generation of candidate for incremental and interactive frequent pattern mining. It needs two times of scan dataset than existing. In incremental and interactive frequent pattern mining the method is very efficient and scalable. IncrShrFP has the build once mine many properties. C. Nawapornanan, et.al, [8] proposed an efficient algorithm for mining share-frequent item sets from Bit

Table that extracted knowledge from a database. The algorithm finds all frequent itemsets by level-wise candidate generation from a Bit Table using heuristics and testing for better solutions.

Systolic tree-based structure [6] mined the frequent item sets. This architecture performed on higher throughput. It provides detailed analysis and required systolic tree-based architecture performance. The result showed that tree-based architecture work faster than existing for mining long frequent patterns. J. Han, et.al, [10] presented FP-tree structure stored compressed information about frequent patterns with a FP-tree based pattern growth method which avoids repeated database scans and a large number of candidate generation sets. The result showed that the FP-growth method is efficient for mining frequent patterns and faster than the Apriori algorithm.

R. Agarwal, et.al [11], enhanced lexicographic tree to generate frequent itemset with different strategies such as breadth first and depth first search. It provided edibility of picking the correct strategy during the tree generation phase. Also proposed parallelization of Tree Projection algorithm has reduced communication required by a large factor compared to the parallel version of Apriori algorithm and more advantages over the parallel implementation of the existing algorithm. In [13] generated an improved Apriori algorithm to reduce the memory space when performed large number of transaction to be scan. It was less time consumed by 67.87%. Hence it was far more efficient than original Apriori. It uses the approach of parallel algorithm and clustering method.

III. PROPOSED WORK

We introduce an advance technique for finding frequent patterns that is ASPMS algorithm. This algorithm finds frequent patterns among sensors in WSN. It can capture important knowledge from the stream contents with single scan of database. The compress feature of ASPMS will show the utilization of less memory than existing algorithm.

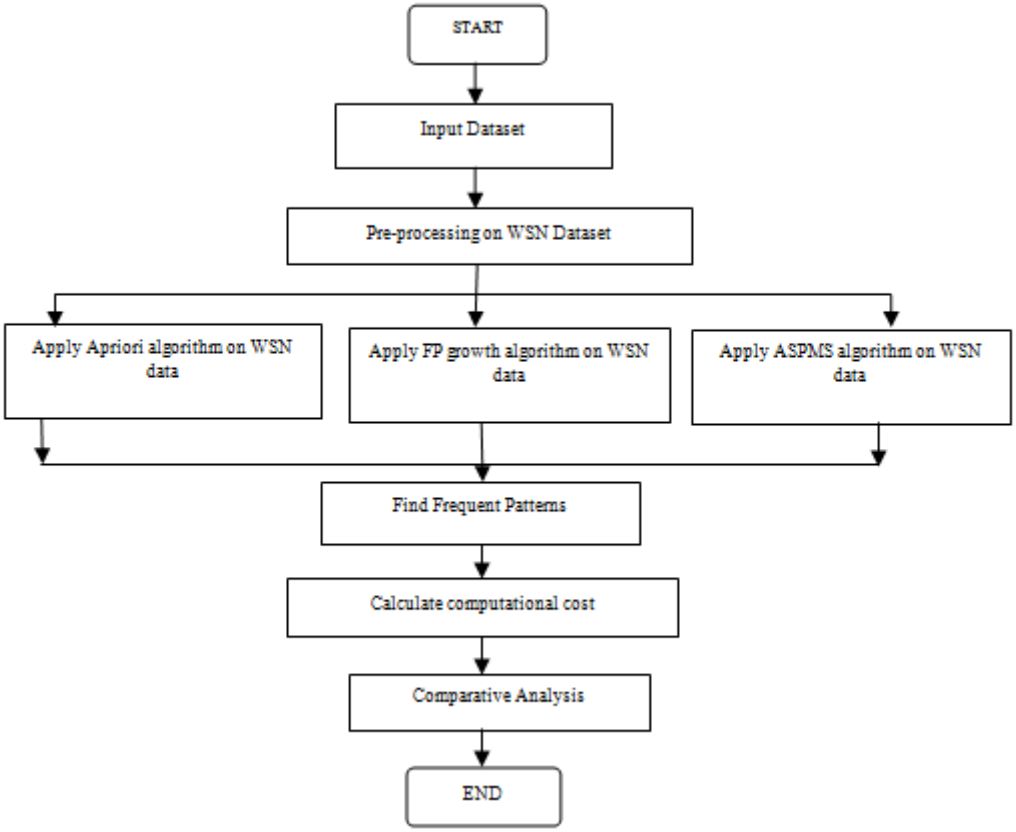

Figure 1.Flow of Proposed System

Proposed system will have WSN dataset as an input data and pre-processing on that dataset. WSN data is the data of truncations that provided by sensors in network. We have use WSN data to find the frequent patterns having approximately 1 lac transaction records that get from sensors. For finding the frequent pattern apply Apriori, FP Growth and ASPMS algorithm on WSN dataset. ASPMS capture important knowledge from stream contents. Calculate computational cost of algorithms for improving quality services of WSNs. We compare both three algorithms for comparative analysis on efficient result of algorithms.

A. ASPMS Algorithm

Different techniques applied in frequent pattern mining to find candidates and frequent patterns generated. In frequent pattern mining techniques, there are two problems. First is that many times scanned the database and second is more complex candidate generation process for finding frequent patterns. We propose a technique ASPMS Algorithm [14] with ASPS-tree. ASPS-tree and algorithm is used to generate associated patterns.

ASPMS algorithm can extract associated sensor patterns in the current window with FP growth like pattern-growth method after getting useful information from the ASPS-Tree. ASPMS perform single scan of database and more flexible

for transactions. This algorithm can extract useful information from the stream contents for the current window of the sensor in a batch-by-batch manner. ASPS-tree is based on sliding window that associated sensors pattern for WSN. The nodes of an ASPS tree in a sensor appearance order and then restructure the tree in a frequency-descending order. Then finally compress the tree by merging the same support sensor node in a single node in each branch of the tree.

By using BSM, it compresses the same frequency nodes into single node. BSM is keep the frequent items in tree after restructure phase. In this method if path is not sorted according to new insertion list order, it is removed from the tree, deleted non-frequent items, sorted according to new insertion list order into a temporary array and then again inserted into tree in order.

B. Working of ASPMS algorithm

Step 1: Original database is dividing into equal size of windows, then each window contains equal number of batches and each batch contains equal no of transactions.

Step 2: Consider first window and form a tree according to their sequence. Complete tree for all transactions in first windows.

Step 3: Before transactions of next windows are inserted batch by batch, tree needs to be compressed and restructured.

Step 4: Sort the SO-List in support descending order and restructure tree according to SO-list using merge sort.

Step 5: If two nodes of same branch have same support count, merge them into single node using BSM.

Step 6: Complete ASPS tree after inserting next window transactions batch by batch and compressed the tree.

Step 7: Extract associated frequent sensor patterns from the ASPS tree.

This technique has two phases: Tree construction and mining associated sensor patterns. Tree construction phase is further divided into two sub phases; insertion phase and restructuring with compressing phase.

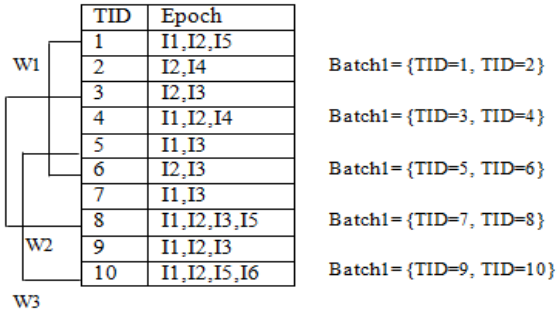

Figure 2.Example of Original Database of sensor data streams.

From above example the original database is divided into 3 equal sizes of windows. Windows 1 contain batch 1, batch 2 and batch 3. Batch 1 contains TID1& TID2. According to their sequence form a tree with all transaction in window. Before transactions of windows 2 are inserted batch by batch. Tree needs to be compressed and restructured, so sort the SO-List in descending order and restructure tree according to SO-list using merge sort. Use BSM for merge same nodes into single node, if nodes of same branch have same support count. Complete ASPS tree after inserting third window transactions batch by batch and compressed the tree. We expect from work of ASPS tree that will show the efficient result and find frequent patterns of transactions after reconstruct and compress a tree.

The features of ASPMS algorithm is cost efficient because of less iteration to find frequent patterns in the respective subsets. Therefore it reduces more computational cost. Because of reduce computational cost effective nature it is time complexion**.** It will require less memory with single scan of database. ASPMS is more flexible for addition and deletion of transactions.

IV. CONCLUSIONS

In this work, we have proposed ASPMS algorithm with ASPS-tree can extracts more useful knowledge from sensors data. It can find efficient frequent pattern with one scan database as compare to existing algorithms. The expected outcomes from this algorithm are to retrieval of frequent pattern find as output using ASPMS. It will show the efficient result than Apriori and FP Growth in comparative analysis with parameter like memory-scan, computation time, efficiency and storage structure.

REFERENCES

[1] Agrawal.R. and Srikant.R, "Fast algorithms for Mining Association Rules", 20th International conference on Very Large Data Base, pp. 487–499, May1994.

[2] Md. Mamunur Rashid, Iqbal Gondal and Joarder Kamruzzaman, "Share-Frequent Sensor Patterns Mining from Wireless Sensor Network Data", IEEE Transaction Parallel Distribution System , 2014.

[3] Y.C.Li, J.S. Yeh and C.C.Chang, "A fast algorithm for miningshare-frequent itemsets", in Proceeding. AP Web, pp. 417-428, 2005.

[4] Jun TAN, Yingyong BU and Bo YANG, "An Efficient Frequent Pattern Mining Algorithm", Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009.

[5] K.M. Yu, J. Zhou and W.C. Hsiao, "Load Balancing Approach Parallel Algorithm for Frequent Pattern Mining", Proceedings PaCT, pp.623-631, 2007.

[6] S. Sun and J. Zambreno, "Design and analysis of a reconfigurable platform for frequent pattern mining," IEEE Transaction Parallel Distribution System, Vol.22, no. 9, pp. 1497, 2011.

[7] Anjan Das, "An Enhanced Data Reduction Mechanism to Gather Data for Mining Sensor Association Rules", IEEE Transaction Parallel Distribution System, pp. 120-136, 2011.

[8] C. Nawapornanan and V. Boonjing, "An efficient algorithm or mining complete share-frequent item sets using Bit Table and heuristics", Proceeding of ICMLC, Vol.1, pp.96-101, 2012.

[9] A. F. Chowdhury, T. K. Syed and J. Byeong-Soo, "An Efficient Method for Incremental Mining of Share-Frequent Patterns" in International Asia-Pacific Web Conference, 147-153, 2010.

[10] Jiawei Han, Jia Pei, Yiwen Yin and Runying Mao, "Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach", Data Mining and Knowledge Discovery, 8, 53–87, 2004.

[11] Ramesh C. Agarwal, Charu C. Aggarwal and V.Prasad, "A Tree Projection Algorithm For Generation of Frequent Itemsets", Vol. 61, Issue 3, pp 350-371, March 2001.

[12] M.M. Rashid, I.Gondal and J.Kamruzzaman, "Mining Associated Patterns from Wireless Sensor Networks", IEEE Transaction on Computers, August 2014.

[13] Akshita Bhandari, Ashutosh Gupta and Debasis Das, "Improvised Apriori algorithm using frequent pattern tree for real time applications in data mining", Vol. 46, Pages 644–651, 2015.

[14] Shamila Nasreen, Muhammad Awais Azam, Khurram Shehzad, Usman Naeem and Mustansar Ali Ghazanfar, "Frequent Pattern Mining Algorithms for Finding Associated Frequent Patterns for Data Streams: A Survey", EUSPN, pp.109 – 116 -2014.

