International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

Efficient implementation of data mining techniques in wireless sensor networks

M.Agnisha & S.Srilakshmi
Asst.Professor
St.Martin's Engineering College Affiliated to JNTUH.

Abstract:

The wireless sensor nodes are getting smaller, but Wireless Sensor Networks (WSNs) are getting larger with the technological developments, currently containing thousands of nodes and possibly millions of nodes in the future. Sensor Networks (SNs) produces huge amount of data which offer promising prospect for the application of data analysis techniques to extract useful information for end users. This paper deals with the efficient data mining techniques in wireless sensor networks

Keywords *Data Mining Techniques, Wireless sensor networks, classification, clustering*

1. Introduction: Over the past few years, wireless sensor networks (WSN) have emerged as one of the most exciting fields in Computer Science research. A Wireless Sensor Network (WSN) is a set of sensors that are integrated with a physical environment. These sensors are small in size, and capable of sensing physical phenomenon and processing them. The main aim of deploying applications based on WSNs is to make use of the data sensed by the sensors to raise the real-time decisions. The main limitations of WSNs are characteristics of sensor nodes and nature of sensor data generated by networks. Due to these limitations, traditional data mining techniques are not suitable to WSNs. As the data generated by WSNs is highly resource-constrained, huge in volume, fast changing, it is very challenging to design suitable data mining techniques for WSNs**.**

2.Applications of Wireless Sensor Networks :- The Applications for WSNs involve tracking, Monitoring and Controlling. WSNs are mainly utilized for habitat monitoring, object tracking, nuclear reactor control, fire detection and traffic monitoring. There are lots of applications in wireless sensor networks [1]

•Process Management: Area monitoring is very common using WSNs. In area monitoring, the WSN is deployed spanning a region where some phenomenon is usually to be monitored. A military example may be the use of sensors detect enemy intrusion; a civilian example would be the geo-fencing of gas. Area monitoring is most important part.

•Healthcare monitoring: The medical applications might be of two sorts: 9Wearable 9Implanted Wearable devices are applied to the body surface of the human or maybe at close proximity from the user. The implantable medical devices are the ones that are inserted inside your body. It is used to monitoring of ill patients in hospitals and also at home.

•Environmental/Earth sensing: There are numerous applications in monitoring environmental parameters samples of which are given below. They share any additional challenges of harsh environments and reduced power supply.

•Polluting of the environment monitoring: Wireless sensor networks have been deployed in lots of cities (Stockholm, London and Brisbane) to monitor the power of dangerous gases for citizens.

•Forest fire detection: A network of Sensor Nodes is usually positioned in a forest to detect every time a fire has begun. The nodes are usually with sensors to measure temperature, humidity and gases which are produced by fire within the trees or vegetation. The first detection is necessary to get a successful action of the fire fighters; As a result of Wireless as Sensor Networks, the fire brigade are able to know when a fire begins you bet it can be spreading.

•Water quality monitoring: Water quality monitoring involves analysing water properties in dams, rivers,lakes & oceans, and also underground water reserves. All the applications which are mentioned above are reliable and real-

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

time monitoring is the essential requirements. All the application generate huge volume of data which are fast changing and heterogeneous in nature. All the data is collected and filtered into useful information or pattern by using data mining techniques. As in upcoming years the whole world will be ruled by wireless sensors, so it will be very crucial to choose a correct data mining technique which is a big challenge in WSNs.

3.DATA MINING TECHNIQUES :-The data mining is elucidated as heart of enlightenment discovery process. It can be briefed as the process of gathering facts from different area and refining it into appropriate knowledge. It can also be defined as the evocation of hidden predictive knowledge from a huge database. Technically, from
a large relational databases a search is performed among several field areas to obtain a useful pattern which can be used in future. The motive of using data mining is to help companies to focus on the most important data in their data warehouses. The main idea of data mining technique is to extract data from large dataset and convert into some useful pattern for future use. Data mining is divided into two models, descriptive and predictive. All Data mining techniques fall under these two categories [4].
•Predictive model - The primary goal of using this exemplary is that we can predict the future result than the current situation. It falls under the supervised learning and the predicted output can be numeric as well as in categorized form, as it always predicts the target value.
•Descriptive model - This method is generally used to generate correlation, frequencies, cross tabulation etc.
It is used to discover regularities in the data and uncover patterns. From bulk of data, a search is performed for finding interesting subgroup patterns.

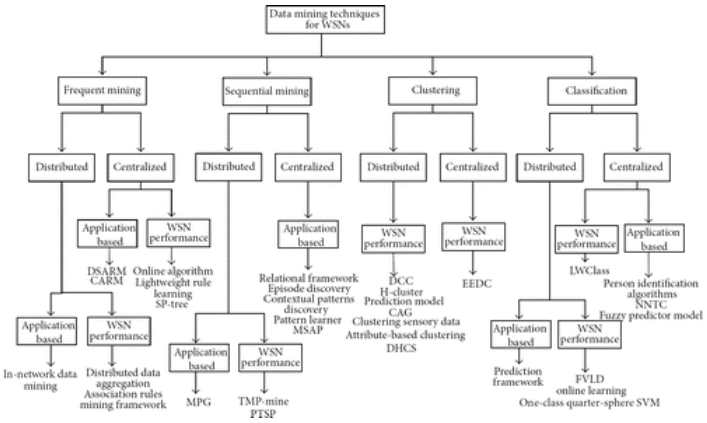

Figure 1 Taxonomy of data mining techniques for sensor networks.

4.State of the Art of Data Mining Techniques for WSNs

In this section, data mining techniques designed for WSNs are classified using the taxonomy framework presented and the characteristics and performance analysis of each technique is discussed.

Frequent Pattern Mining

In this section, we review some of the works that have been proposed for mining frequent patterns from WSNs data. Frequent pattern mining is used to find the group of variables that co-occur frequently in the data-set. The aim is to find the most interesting relations between variables. Traditional frequent pattern mining algorithms are the CPU and the I/O intensive, making it very expensive to mine dynamic nature of WSN data. Unlike the mining static database, dynamic nature of WSNs data led to the study of online mining of frequent itemset. As a result, traditional frequent pattern mining algorithms are modified according to nature of WSNs data.

The basic frequent pattern mining technique is association rule mining technique. The first known association rule mining algorithm is *Apriori*. It is based on level-wise candidate generation and test methodology by making several scans over database. In each iteration, the patterns found to be frequent are used to

generate possible frequent patterns (the candidates) to be counted in the next iteration. Therefore, the *Apriori* technique finds the frequent patterns of length k from the set of already generated candidate patterns of length $k-1$

. In the subsequent step, the association rules are generated by computing the support and confidence of each frequent item in given database D which is defined as follows:

$$\text{Support}(A)=\text{Sup}(A)D, \quad (1)$$

where Sup(A)

is the number of occurrence of A in database D. Consider the following:

$$\text{Confidence}(A\longrightarrow B)=\text{Sup}(A\cup B)\text{Sup}(A). \quad (2)$$

This is impractical in the context of sensor networks as it implies that all data has to be stored somewhere. However, recently, there has been a growing amount of work on discovering frequent item-sets from a data *stream* of transactions such that every transaction is considered only once and can be deleted afterwards.

The other basic approach from mining association rule is FP-growth which can discover frequent patterns by reducing the database scans by two and eliminating the requirement of candidate generation as compared with Apriori. With the first database scan, the algorithm finds the set of distinct items with respective support count (i.e., frequency) in the database. Then, with the second database, scan the algorithm summarizes the database in the form of a frequency-descending tree (i.e., the FP-tree). The complete set of frequent patterns is, then, mined from the FP-tree by recursively applying a divide-and-conquer-based pattern growth approach, called the FP-growth algorithm, without additional database scan. The highly compact FP-tree structure introduced a new wing of research in mining frequent patterns. However, the static nature of the FP-tree and two database scans still limit its applicability to frequent pattern mining over a WSNs data. Recently, several centralized and distributed solutions have been proposed with the aim to maximize the WSNs' performance and maximize the application-based performance by applying Apriori-like and FP-growth methods over WSNs data.

5.CONCLUSION: The key challenge of evolving intrusion detection system in WSN is to identify attacks with high accuracy, and satisfied the required constraints and challenges, to prolong the lifetime of the entire network. This aims could be attained from several ways. Firstly paying much more attention to detection techniques used for attacks detection is characterized by efficiency and ability. Secondly, reconstructing detection mechanism with a distributed manner, to reducing the communication overhead. This paper has compared and evaluated the newest anomaly detection intrusion techniques used in wireless sensor network, to improve the efficient technique for IDS in WSN. According to the results, it is highly recommended to use the data mining techniques to detect effectively the intrusions and attacks in WSN. The decision of choosing efficient IDS is a compromise between technique employed and performance metrics. However, many issues are still open and need further research efforts such as hierarchical clustering patterns, using machine learning in resource management problem of wireless sensor networks, developing a classifier that is trained well with network patterns, selecting and preprocessing an appropriate dataset. In addition, taking smart strategies into account such as compressing the input dataset, narrowing the scale of attributes set and simplifying the procedure of analysis and decision could make lots of progress for IDS to satisfy the requirement constraint of WSN without losing the security and reliability.

6.Acknowledgements: This work was supported by my family,colleagues friends and my college .

7.References:

[1].Akyildiz,I.F.,Su,W.,Sankarasubramaniam,Y.,Cayirci,E.,2002. Wirelesssensor networks: asurvey. Comput Networks 38,393422.

[2]Lopez J, Zhou J. Overview of wireless sensor network security. In: Wireless sensor network security. IOS Press, incorporated; May 2008. p. 1–21.

[3]Perrig A, et al. SPINS: security protocols for sensor networks .Presented at the 17th ACM international conference on mobile computing and networks, 2001.

[4] Kumar, D.Monitoring forest cover changes using remote sensing and GIS: a global prospectiveResearch Journal of Environmental Sciences20115105123 Google Scholar, Crossref

[5] A. D. Wood and J. A. Stankovic,(2002) "Denial of service in sensor networks",Computer, 35(10):54–62, 2002.

[6] Jianfa Xia, Zhenzhou Tang, Xiaoqiu Shi, Lei Fan,Huaizhong Li " An environment monitoring systemfor precise agriculture based on wireless sensornetworks" 978-0- 7695-4610- 0/11 $26.00 © 2011IEEE