

APPLICATION OF DATA MINING TECHNIQUES IN WIRELESS SENSOR NETWORKS: A REVIEW

S. Shylaja¹

1Assistant Professor & Research Scholar,
Department of Computer Science,
Rathinam College of Arts & Science (Autonomous)
Coimbatore-641021, Tamil Nadu, India,
shylaja.cs@rathinamcollege.com

R. Muralidharan²

2 Vice Principal & Head,
Department of Computer Science,
Rathinam College of Arts & Science (Autonomous),
Coimbatore-641021, Tamil Nadu, India,
hod.cs@rathinamcollege.com

Abstract - Wireless Sensor Network is one of the most important technologies, which is being used in computer science. It consists of smart sensor nodes which can sense and communicate in a wireless fashion in a defined network and it get the data to sink. The sink is connected to gateway and collected data's are sent to the base station. Sensors have many constraints like limited resources, energy, memory, computation power etc. Data Mining is used to evaluate data from several dimensions and is the process of finding pattern in a large relational database. The purpose of using data mining is to increase the efficiency of the sensor nodes. The main objective of this review focused on distinct data mining techniques which have been adopted by WSN showing their advantages and disadvantages with various types of applications.

Keywords - Data Mining Techniques, Wireless sensor networks, Centralized approach, Distributed approach

1 INTRODUCTION

A. Wireless Sensor Networks: Advantages and Disadvantages

In Today's scenario wireless sensor network has become very essential in our daily uses, it is considered as one of the most important technology for the twenty-first century without WSN our work would have been very difficult. WSN can reach into those areas where human cannot even think to reach like disaster areas, space, sea etc. WSN has earned terrific attention from both industry and academic all over the world. Wireless sensor network consist of thousands of low cost nodes, low-power, and multifunctional sensor that are deployed in a region of interested area which run on battery. The sensor nodes generally are small in size each node has one or more embedded microcontroller i.e. CPU or DSP chip set. These nodes have sensing capabilities along with communication capabilities. These nodes interact over a short area through a wireless medium which are organized into ad-hock network [7].

Advantages and Disadvantages of WSN

Why people love wireless sensor networks might be summarized as the following [12]:

- Network setups can be carried out without fixed infrastructure.
- Suitable for the non-reachable places such as over the sea, mountains, rural areas or deep forests.
- Flexible if there is random situation when additional workstation is needed.
- Implementation pricing is cheap.
- It avoids plenty of wiring.
- It might accommodate new devices at any time.
- It's flexible to undergo physical partitions.
- It can be accessed by using a centralized monitor.

The disadvantages of wireless sensor networks can be summarized as follows [12]:

- Less secure because hackers can enter the access point and obtain all the information.
- Lower speed as compared to a wired network.
- More complicated to configure compared to a wired network.
- Easily troubled by surroundings (walls, microwave, large distances due to signal attenuation, etc.).
- It is easy for hackers to hack it we couldn't control propagation of waves.
- Comparatively low speed of communication.
- Gets distracted by various elements like Blue-tooth.
- Still Costly (most importantly)

B. 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 [12]

- 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:
 - ✓ Wearable
 - ✓ Implanted

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-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.

The pattern of paper is as follows: Section2 explains about data mining technique. Section 3 explains about WSN in brief. Section 3 deals with the literature survey and section 4 describe the conclusion of the whole paper.

II 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.

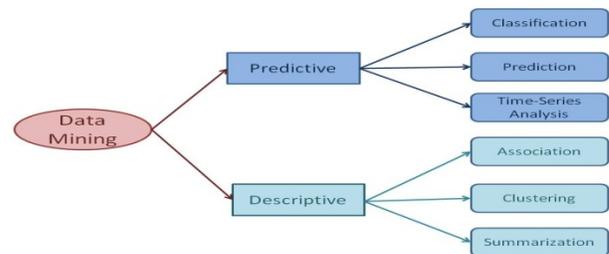


Figure 1 : Data Mining Models

The main steps are for knowledge discovery stage is as follows:

1. **Data Cleaning:** It is the first step in which inconsistent data and noises are removed.
2. **Data Integration:** It is the second step where the combination of multiple data source is done.
3. **Data Selection:** It is the third step where necessary data are searched from the database to reach the goal.
4. **Data Transformation:** It is also called as data consolidation. In this stage the selected e data are transformed into useful pattern for mining
5. **Data mining:** It is the most crucial stage in which the data mining techniques are applied on transformed data so that the tendency of the extracted pattern can be determined.
6. **Pattern Evaluation:** It is the sixth step in which only the important pattern which holding accurate information are chosen.

Knowledge Presentation: It is the final stage in which the mined data are provided to the user using virtualization and knowledge representation techniques

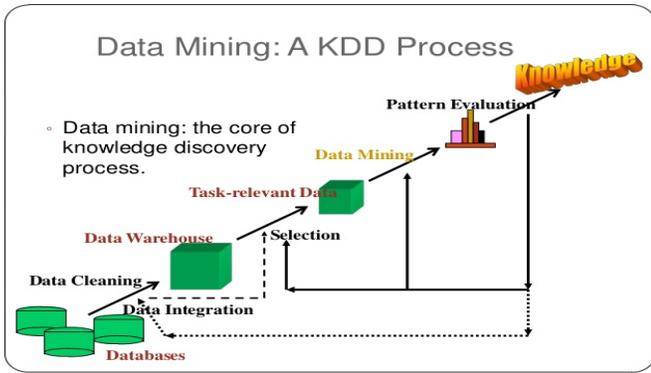


Figure 2: Stages for Knowledge Discovery process

III DATA MINING IN WIRELESS SENSOR NETWORKS

Data mining in sensor network is the method of selecting application oriented standards and patterns with acceptable accuracy from a continuous, fast and probably non ended flow of data streams from sensor networks. In this case, all data cannot be stored and must be processed quickly, Data mining method has to be sufficiently fast to process high speed arriving data. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use [15]

There are four types of data mining techniques which are commonly used in WSNs which uses both centralized and distributed approach that is given below

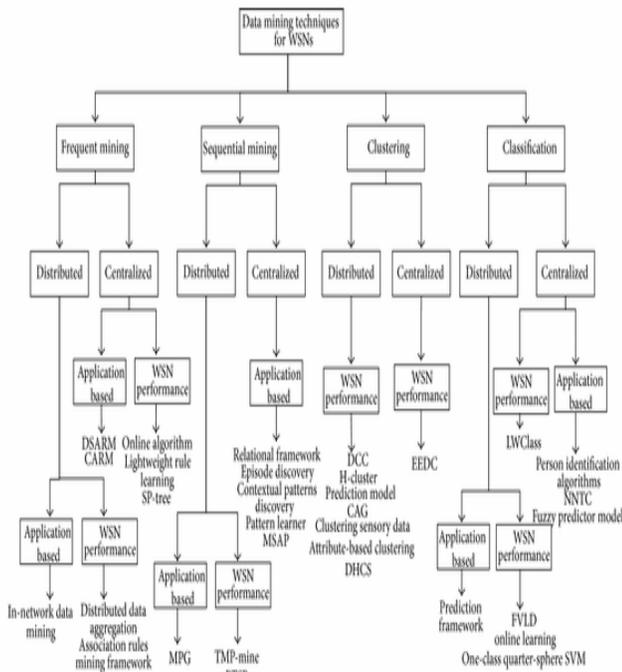


Figure 3: List of Data mining techniques in Wireless Sensor Networks

There are three essential elements for data mining without these nothing can be achieved, They are as follows;

1. **Data:** are any facts, numbers, or text that can be processed by a computer and these growing amounts of data in different formats and different databases are stored. The data which defines other data is called met data such as design of logical database.
2. **Information:** Any data which provide information in form of pattern or relationship.
3. **Knowledge:** The patterns which are formed or gained are converted into knowledge for future use

IV WIRELESS SENSOR NETWORKS

A wireless sensor network (WSN), it is a wireless network which dwell of partially disperse autonomous equipment's that uses sensors to monitor environmental and real condition. A WSN is a collection of large sensor nodes which can be in hundreds or even thousands of small, cheap nodes which are deployed into a network at certain location. These sensors have the capability to sense, process and communicate to its peer in order to work together in a cooperative manner. A sensor network consist of thousands of battery which sense huge bulk of data and send it back to the Sink through gateway for processing purpose by using different data mining techniques. Each node consists of processing capability in form of microcontrollers that is CPU or DSP chip sets. WSNs have different endowment like denser level of node deployment, sever energy, computation power, data generated by sensor nodes are bulk in size, fast changing and have very less resource constraints [3]. In figure shown below consist of thousands of sensor node which are arranged in a desired location, the jobs of sensor nodes are to collect bulk of data and transfer it to number of gateway. They get the data to a sink and the sink is wired to a gateway over a network, all the data are transfer to the base station where useful data are sorted for future work. The nodes communicate with each other in a wireless manner which follows ad-hoc network. At present WSNs consist of thousands of nodes which in future may contain more than millions of nodes. In order to solve huge volume of data we need to solve all the major issues of the sensors like processing capability, limitation in memory, etc.

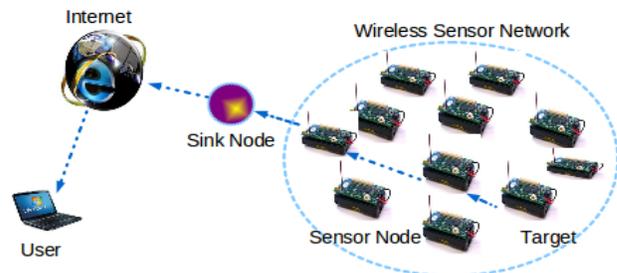


Figure 4: Wireless Sensor networks

Other issues these nodes are, in future we have to deal with heterogeneous data like sound, smell, taste, image, audio, video, etc. In this kind of data, significant amount of energy is consumed. There are basically two types of approaches in WSNs which are mentioned below:

Table 1 Difference between Distributed and Centralized Sensor

Application	Distributed	Centralized
Energy	Limited	No Constraints
Data Length	Un Bounded	Bounded
Response time	Real Time	Non Real Time
Update Speed	High	low
Number of Passes	Single	Multi task
Data Flow	Continuous	Stationary
Data Type	Dynamic	Static

A. Centralized Approach

Generally the data mining technique uses centralized approach. In this approach data are collected directly at a central location which is not protected or restricted by any computational resources. This approach involves the selection of cluster head by the base station itself. It is done on the basis of amount of residual energy and distance of sensor nodes from base station. In today’s scenario the data which are generated are very fast developing and are huge in volume as the memory storage is limited. The main problem in this approach is, all the data collected are transferred to the central node and if the central node fails then the whole network will fail. This approach is mostly inappropriate due to scalability issue like limited communication bandwidth, power supply and less storage space.

B. Distributed Approach

This approach is better than centralized approach as the data is not collected at a central node, if one node fails it will not have any effect on any other nodes. In this method the resources are need to be self-organized. The reliability of sensor nodes can be improved by the inducement of distributed architecture. This approach helps in saving energy up-to a great extent. This technique is used in order to control the issues like:-[7]

- Sensor nodes are prone to failure
- Better collection of data at sensor nodes
- To provide backup with nodes if failure occurs in the central node.

V RELATED WORKS

Zahra haihashemi and mihail [1] proposed the A multidimensional Time Series similarity measure with application to eldercare monitoring. Here, data mining techniques have been applied to sensor data in a wide range of application domains such as Health monitoring systems, manufacturing processes, Intrusion detection, database management and others. Many data mining techniques are based on computing the similarity between two sensor data patterns. Then our author described a novel method for computing the similarity of two multi attribute time based on

temporal version of smith waterman (TSW-Temporal smith waterman) is a well-known bio informatics algorithm. This method was used to sensor data from eldercare applications for early illness detection. And these types of methods mitigate difficulties related to data uncertainty and aggregation that often arise when processing sensor data. Finally, The WTSW (Window based temporal smith waterman) could be potentially slow for eldercare applications. So they propose a genetic version of it GATSW (Genetic algorithm with TSW) and they tested our algorithm on multiple datasets.

Marjorie skubic et.al [2] presents an example of unobtrusive, continuous monitoring in the home for the purpose of accessing early health changes. Sensors embedded in the environment capture behaviors and activity patterns. A one dimensional (1-D) alter algorithm was used to generate health alerts to clinicians in a senior housing facility. Here, there were four classification approaches that use multi sensor data. Results are shown using embedded sensor data and health alert ratings collected on 21 sensors over nine months. In this paper they were used four classifier techniques: Fuzzy pattern tree (FPT), Fuzzy K-Nearest Neighbor (FKNN), Neural Networks (NN), and Support Vector Machine. Those multi-dimensional classifiers performed significantly better than the 1-D algorithm, with the best 6D performance at 86% compared to the 39% at 1D. And the 6D classifier based only on the Domain Knowledge performed better than the best 6D classifier using supervised machine learning.

Azhar Mahmood [3] has proposed the survey for Mining Data Generated by Sensor Network. The Sensor Network’s (SNs) produces huge accounts of data which offer promising prospect for the application of data analysis techniques to extract useful information for the end user. The DM community had observed that extracting knowledge from SNs Data through two methods. Network side processing technique, that requires real time analysis methodology and to handle dynamic data streams or events. Centralized processing through high end computing is required for generating off-line predictive insights which facilitates real-time analysis. From analysis, it was observed that the techniques intended for mining sensor data at N/W side are helpful for taking real time decision as well as serve as prerequisite for development of effective mechanism for Data storage, retrieval, query and transaction processing at the Central Side.

D.J.Dechnene [4] had conducted a survey of clustering algorithm for Wireless Sensor Networks. They described the improvements to be made in the clustering algorithm for wireless sensor networks. They suggested the following three types of schemes **Heuristic Scheme, Weighted Schemes, Grid Schemes**

Marsinisso Saoudi [5] had described the Data mining Techniques that are applied to Wireless Sensor Networks for Early Forest Fire Detection. They proposed a new approach for forest fire detection based on the integration of data mining techniques into sensor nodes. The idea is to use CLUSTERED WSN.

VI CONCLUSION

In this paper, we have explained different data mining method which has been adopted for wireless sensor network which uses different classification and evaluation approaches. We inspected that each data mining technique mentioned here works on either centralized or distributed approach are discussed. However the main consideration of the approach was information extraction and analysis on data with respect to WSNs, We have analyzed and briefly described both advantages and disadvantages of each algorithm. The entire algorithm which has been discussed above is capable of solving different types of issues related to WSNs types and applications. The major challenges will be related to hardware like compressing, storing and filtering of huge volume of data generated by sensor nodes. Also in future we have to deal with heterogeneous kinds of data like images, sounds, smell, location and etc. The main motive to present this paper is summarize, analyze and extract knowledge from data into useful ideas for better decision making.

REFERENCES

[1] Zheng and Jamalipour, A., "A wireless sensor networks: A networking prospective", John Wiley & sons, 2009.

[2] Zhao, X. and Sayed, A. H. (2012). Performance limits for distributed estimation over lms adaptive networks. *Signal Processing, IEEE Transactions on*, 60(10):5107–5124.

[3] A.R Raut and S.P Khandait, "Review on Data Mining Techniques in wireless sensornetworks" *Electronics and communication systems (ICECS) 2015, 2nd International conference on Coimbatore 2015*, PP 390-394.

[4] Pradnya P. Soundwale, "Overview of predictive and descriptive Data Mining techniques", *IJARCSSE, Vol-5, 2015*, PP 262-265.

[5] Bouyer, A. Hatamlou, A and Masdari, M "A new approach for decreasing energy in wireless sensor networks with hybrid LEACH protocol and fuzzy c-means algorithm" *International journal of communication network and distributed system*, 15(4), 2015, PP 400-412

[6] Shrikant D. Dhamdhare, Shanthi K, Guru "Robust Data collection in wireless sensor networks with mobile sink", *International journal of Computer Science and Information technologies*, vol 5 (4), 2014, PP 4999-5002.

[7] Aditya kumar naik, Rajendra kumar, Dwivedi "A Review on use of Data Mining methods in wireless sensor networks" *ISSN (Print) 2393-8374, Online: 2394-0697, Vol 3, Issues 12, 2016, International Journal of current engineering and scientific research (IJCESR)*.

[8] Massinissa saoudi, Ahcene Bounceur, Reinhardt Euler: Data Mining techniques applied to wireless sensor networks for early forest fire detection" *ISBN 978-1-4503-4063-2. ICC'16, March 2016, Cambridge, United Kingdom ACM*.

[9] Mirjana Maksimovic, Vladimir Vujovic "Comparative analysis of Data Mining techniques applied wireless sensor networks data for fire detection" *JIIT3 (2013)2:65-77, Journal of information technology and application, December 2013*.

[10] Mohammed Alsheikh, Shaowei Zin, Dusit Niyato and Hwee-Pink Tan "Machine learning in Wireless sensor network: Algorithm, Strategies and Applications, arXiv: 1405.4463v2 [cs.NI] 19 March 2015.

[11] Shubhie agarwal, Seema Matrey, Pooma Rana, Pankaj Singh Yadav "A reasonable exploration of Data Mining techniques in WSN: International of journal of Computer applications (0975-8887) vol 151- No 10, Oct 2016.

[12] Prashant Tiwari, Varren prakash saxena, Raj Gauravmishra, Devendra Bhavsar, "Wireless sensor network: Introduction, advantages, applications and research challenges" *HCTL open international journal of technology innovation and research (IJTIR) Vol 14, April 2015, ISBN 978-1-6-2951-946-3*.

[13] Marjorie skubic, Rainer Dane Guevara and Marilyn rantz "Automated health alerts using in home sensor data for embedded health assessment" *IEEE journal of transactional engineering in health and medicine*. Received 29 Oct

2014, Revised 28 Dec 2014, Accepted 6 mar 2015, Date of publication 10 Apr 2015. Date of current version 22 Apr 2015.

[14] Azhar Mahmood, Ice Shi and Shaheer Khatoon "Mining Data Generated by Sensor Network, A survey" *Information Technology Journal II : 1532–1543, 2012 ISSN 1812-5638*

[15] Sherin Mathew, —A Novel Approach for Sensory Data Collection in Wireless Sensor Networks with Mobile Sinks□, *International Journal of Engineering Science and Innovative Technology (IJESIT)*, Volume 2, Issue 2, March 2013

[16] Shrikant D. Dhamdhare, —Robust Data Collection in Wireless Sensor Networks With Mobile Sinks□, *International Journal of Computer Science and Information Technologies*, Vol. 5 (4), 2014

[17] S.Punitha Shankari -Energy Saving In Wireless Sensor Network Using Burst Algorithm□, *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, Volume 5, Issue 1 (Jan. Feb. 2013)

[18] V.Karthik, -Region Based Scheduling With Multiple Mobile Robots for Data Collection Strategies in Wireless Sensor Networks□, *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 2, Issue 7, July 2013

[19] Fan Bai, - A Novel Information Acquisition Technique for Mobile Assisted Wireless Sensor Networks□, *IEEE Transactions on Vehicular Technology*, Vol. 61, No. 4, May 2012

[20] MiaoZhao, —Bounded Relay Hop Mobile Data Gathering in Wireless Sensor Networks□, *IEEE Transactions On Computers*, Vol. 61, No. 2, February 2012

[21] S. Nithyakalyani, —Data Relay Clustering Algorithm for Wireless Sensor Networks: A Data Mining Approach□, *Journal of Computer Science 8 (8): 1281-1284, 2012*

[22] Khushboo Sharma, -Nearest Neighbour Classification for Wireless Sensor Network Data□, *International Journal of Computer Trends and Technology*, volume2, Issue-2, 2011

[23] AshuGupta and R.Annie Uthra, Cluster Based Approximate Data Collection in WirelessSensor Network, *IJECSE, Volume2, Number 2*

[24] Jyoti Rajput, A Survey on Secure Data Aggregation in Wireless Sensor Network□, *IJARCSSE, Volume 4, Issue 5, May 2014*

[25] LaxmiChoudhary, —Challenges for Data Mining□, *IJREAS, Volume 2, Issue 2, February 2012*

[26] Mukesh Kumar Jha, T.P Sharma —Secure Data aggregation in Wireless Sensor Network: A Survey, *International Journal of Engineering Science and Technology, ISSN: 0975-5462, Vol. 3 No.3, March-2011*

[27] P. D. Patel, P.B. Lapsiwala, R.V. Kshirsagar —Data Aggregation in Wireless Sensor Network□, *International Journal of Management, IT and Engineering*, vol. 2, Issue 7 July- 2012

AUTHORS PROFILE



National and International

Ms.S.Shylaja is an Assistant Professor & Research Scholar in Rathinam College of Arts & Science, at Affiliated to Bharathiar University, Coimbatore, India. She has more than two years of experience. Her area of interest is Data Mining. And she is published two papers in International journals & presented 8 papers in National and International Conferences.



Dr.R.Muralidharan is working as Vice Principal & Head of the Department of Computer Science in Rathinam College of Arts and Science, Coimbatore, India. He got his Doctoral Degree in Computer Science, And also he has got sixteen years of experience in Academics. He has published papers in 7 Journals and 12 Conferences. - See more at: <http://www.rathinamcollege.edu.in/rmuralidharan/#sthash.uvPCanma.dpuf>