AN OVERVIEW OF OUTLIER DETECTION IN WSN

Ms. Sapna Aggarwal*, Ms. Saima Jan**
*** Assistant Professor,Jind Institute of Engg. & Technology, Jind, Haryana,**
**** Student, Jind Institute of Engg. & Technology, Jind, Haryana.**

Abstract- A wireless sensor network is a network made of huge number of nodes that consist of very less battery life and data processing capabilities. These microelectronics nodes are capable of measuring pressure, temperature and various other environmental conditions. In this paper we will present an algorithm of wireless sensor network having outlier detection system in the same. We will present an energy efficient algorithm with outlier detection system based on trust voting algorithm and sensor rank. Also, we will represent the result after the simulation of our work.

Keywords: Outlier, Deployment, Wireless Sensor Nodes, Clustering, Outlier Identification Labeling**.**

I. INTRODUCTION

Wireless Sensor Network (WSN) is made up of spatially distributed autonomous sensors which can detect changes in the environment like pressure, temperature and other environmental conditions and to considerately pass their data throughout the network to a main location. The more modern networks are bi-directional, enabling control over the activity of the sensors. The growth of wireless sensor networks was aggravated by armed forces applications such as battleground surveillance; these days such networks are also used in many manufacturing and end user application, such as engineering process monitoring and control, machine health monitoring, environment and habitat monitoring, healthcare applications, home automation, and traffic control. The WSN is made up of nodes from a few or hundreds in number, where each node is having one sensor generally (sometimes multiple sensors). Every such type of wireless sensor network node has normally these parts: a radio transceiver with an internal antenna or connection to an external antenna, usually a battery or an embedded form of energy harvesting, a micro-controller, an electronic circuit for interfacing with the sensors and an energy source. A sensor node size may change but it is generally small in size. The cost of sensor nodes is also likewise changeable, having range from hundreds of rupees to thousands, depending on the complexity of the any sensor node that is to be deployed. Size and cost constraints on deployed sensor nodes cause equivalent constraints on resources such as energy, memory, communications bandwidth and computational speed. [6,7] The WSN may have bus, mesh or star topology depending upon the scenario. Also, the propagation methods among the hops of the network can be by routing or by flooding. Outlier detection means a technique of finding for problem in data of any event related to wireless sensor network as per our scenario. These anomalous patterns are frequently called as outliers, noise, discordant observations, exceptions, faults, defects, errors, anomalies, aberrations, damage, contaminants, novelty, peculiarities or surprise in different application domains. In WSNs, outliers can be defined as, “those measurements that significantly deviate from the normal pattern of sensed data” [8].

II. TYPES OF OUTLIER

When the total data is examined as per the central data approach by any central authority then outliers can be recognized properly and can be tackled appropriately at the corresponding station. When type of data is measured the outliers can be divided as local and global outliers:

Local Outliers: considering the point that local outliers are known in wireless sensor network at individual sensor nodes, important task is recognizing techniques reducing communication overhead and maintaining scalability of wireless sensor network with proper determination of outliers. Many event based applications like, surveillance, vehicle following and monitoring can be done by the use of local outlier detection. There are two variations in Local outlier identification in wireless sensor network.

One variation is that historical values are used for determining the wrong or faulty value in the given sensor network. Another option is adding historical reading of their own; where the value of neighbor is taken to determine the value is proper or not i.e. the anomaly is based on the feedback from the neighbor node. When compared with the second approach the first one lags as it doesn’t provide that much accuracy and robustness in the detection of outliers.

Global Outliers: Global outlier are famous for their perspective which is global and they get more attention because they make sure every characteristic of WSN could be completed instead of taking local characteristics. On basis of network architecture classification, we can also identify different type of outliers among the deployed nodes. All the collected data is transmitted to base station node in the centralized architecture. It delay the response time very much and cause a lot of communication overhead. CH (cluster head) gathers the data and recognizes outlier in cluster based approach. It has better response time and energy consumption as compared to the former one.[10]

III. OUTLIER DETECTION

In WSNs, outliers can be defined as, "*those measurements that significantly deviate from the normal pattern of sensed data*" [7]. This definition is based on the fact that in WSN SNs are assigned to monitor the physical world and thus a pattern representing the normal behavior of sensed data may exist. Potential sources of outliers in data collected by WSNs include noise & errors, actual events, and malicious attacks.

Recently, the topic of outlier detection in WSNs has attracted much attention. According to potential sources of outliers as mentioned earlier, the identification of outliers provides data reliability, event reporting, and secure functioning of the network. Here, we exemplify the essence of outlier detection in several real-life applications. [3,4]

- Environmental monitoring
- Habitat monitoring
- Health and medical monitoring
- Industrial monitoring
- Target tracking
- Surveillance monitoring,

IV. METHODS FOR ANALYSIS OF OUTLIER DETECTION TECHNIQUES

Three commonly used methods for evaluation of outlier detection algorithms are

- Detection Rate
- False Alarm Rate
- ROC Curve

The effectiveness of outlier detection techniques can be evaluated quantitatively depending on the number of outliers correctly identified: known as the detection rate and the fraction of normal data incorrectly considered as outliers: known as false alarm rate. The receiver operating characteristic (ROC) curve [3] represented in the form of a 2-D graph is usually used to represent the trade-off between detection rate and false alarm rate. The effectiveness of outlier detection techniques depends on the capability to maintain a high detection rate while keeping the false alarm rate low and large area under ROC curve [8].

V. Challenges of Outlier Detection in WSNs

Extracting useful knowledge from raw data is very difficult job [1]. The complex design and nature of sensor data of wireless sensor network it is difficult to fabricate outlier detection for it. Conventional outlier detection methods are not suitable due to the following reasons.

- **Resource constraints:** The low quality and cheap sensor nodes have severe constraints in resources, like computational capacity, energy and communication bandwidth.
- **High communication cost:** In WSNs, radio communication consume a big portion of energy not the computation in real. Computation cost for a sensor node is much lower than cost of radio communication [2].
- **Distributed streaming data:** Dynamic change can come in streaming data due to different streams. Additionally, the original distribution of data thus streamed cannot be known before receiving. Furthermore, direct computation of probabilities is difficult [5].
- **Dynamic network topology, frequent communication failures, mobility and heterogeneity of nodes:** A sensor network deployed in unattended environments over extended period of time is susceptible to dynamic network topology and frequent communication failures.
- **Large-scale deployment:** Deployed sensor networks can have massive size (up to hundreds or even thousands of SNs). The key challenge of traditional outlier detection techniques is to keep an extraordinary detection rate along with it keeping the rate of false alarm as low as possible. This requires the construction of an accurate normal profile that represents the normal behavior of sensor data [9].
- **Identifying outlier sources:** The sensor network is expected to provide the raw data sensed from the physical world and also detect events occurred in the network. However, it is difficult to identify what has caused an outlier in sensor data due to the resource constraints and dynamic nature of WSNs.

Thus, the main challenge faced by outlier detection techniques for WSNs is to satisfy the mining accuracy requirements while maintaining the resource consumption of WSNs to a minimum. In other words the main question is how to process as much data as possible in a decentralized and online fashion while keeping the communication overhead, memory and computational cost low [10].

V. CONCLUSION & FUTURE SCOPE

In this paper we presented review the basics of wireless sensor network and outlier in the wireless sensor network. We also presented various features of outliers like their types, how they are identified, various sources and degree of outliers.

At last, we represented various challenges in detection of outlier in wireless sensor network. More study can be carried out as review of types of outliers as further classification of local and global outliers. And, various algorithms can be implemented for detection of outliers in the wireless networks.

VI. REFERENCES

[1] Chandola, V., Banerjee, A. and Kumar, V., "Outlier detection: a survey", Technical Report, University of Minnesota, 2007.

[2] S. Rajasegarar, C. Leckie, M. Palaniswami, and J. C. Bezdek, "Distributed anomaly detection in wireless sensor networks", in Communication systems, 2006.

ICCS 2006. 10th IEEE Singapore International Conference on, pp. 1 –5, October 2006.

[3] S. Rajasegarar, J. C. Bezdek, C. Leckie, and M. Palaniswami, "Elliptical anomalies in wireless sensor networks," ACM Trans. Sen. Netw., vol. 6, pp. 7:1–7:28, January 2010.

[4] D. J. Hill, B. S. Minsker, and E. Amir, "Real-time bayesian anomaly detection for environmental sensor data", in proceedings of the 32nd conference of IAHR, 2011.

[5] S. Subramaniam, T. Palpanas, D. Papadopoulos, V. Kalogeraki, and D. Gunopulos, "Online outlier detection in sensor data using non-parametric models", in proceedings of the 32nd international conference on Very large data bases, VLDB '06, pp. 187–198, 2006.

[6] W. Heinzelman, A. Chandrakasan and H. Balakrishnan,"Energy-Efficient Communication Protocol for Wireless Microsensor Networks", , January 2000.

[7] W. Heinzelman, "Application-specific protocol architectures for wireless networks", Ph.D. thesis, Massachusetts Institute of Technology, 2000.

[8] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," IEEE Communications Surveys & Tutorials, vol. 12, no. 2, pp. 159–170, 2010

[9] Z. Yang, N. Meratnia, and P. Havinga, "An online outlier detection technique for wireless sensor networks using unsupervised quarter-sphere support vector machine", in Intelligent Sensors, Sensor Networks and Information. Processing, 2008, ISSNIP 2008. International Conference on, pp. 151 –156, December 2008.

[10] T. Kavitha, A. Chandra, "Wireless networks: a comparison and classification based on outlier detection methods "in CSEA 2012, vol. 4, special issue 1; 2013