The Mechanism of Anomaly Detection in Wireless Sensor Network: An Innovative Approach

Deepak Prakash
Department of Computer Science & Engineering,
Maharishi Markandeshwar University, Mullana, Ambala

ABSTRACT
Wireless Sensor Networks (WSNs) have emerged as one of the most important research areas, large numbers of limited resource sensor nodes are used to monitor the physical environment and report any significant information. Many different anomaly detection systems (ADS) have been proposed in the literature over the years. Now apply an algorithm to increase detection sensitivity. Detection of sensor data irregularities is useful for practical applications as well as for network management, because the patterns found can be used for both decision making in applications and system performance tuning. The problem of irregularities detection is to find those sensory values that deviate significantly from the norm. This problem is especially important in the sensor network setting because it can be used to identify abnormal or interesting events or faulty sensors. Dynamic detection model generated using a combination of different data vectors are required to detect time variant anomalies in WSNs. Decentralized, Individual nodes should perform the anomaly detection independently in the local environment.

The scope of this thesis is to develop and make the ADS scalable and robust against attacks. The communication cost can be reduced if only abnormal sensory values, as opposed to all values, need to be transmitted. It is essential to mine the sensor readings for patterns in real time in order to make intelligent decisions promptly.

General Terms
Wireless sensor network, Anomaly Detection, Security, Algorithms.

Keywords
Wireless Sensor Networks, Anomaly Detection Systems (ADS), Detection Sensitivity, Power Saver, Simulation

1. INTRODUCTION
A Wireless sensor network is composed of tens to thousands of sensor nodes which are densely deployed in a sensor field and have the capability to collect data and route data back to base station. Wireless Sensor Network is used in many application now a days such as detecting and tracking troops, tanks on a battlefield, measuring traffic flow on roads, measuring humidity and other factors in fields, tracking personnel in buildings. Sensor nodes consist of sensing unit, processing unit, and power unit. In wireless sensor network application there are two types of nodes: source node, the node which actually sense and collect data other sink node, the node to which the collected data is sent. The sinks can be part of the network or outside the wireless sensor networks. Usually, there is more number of source nodes than sink nodes.

Data collection from the nodes deployed in the sensing field is one of the main applications of wireless sensor networks. WSNs can be densely distributed over a large area and individual nodes can autonomously communicate and interact with each other over the wireless medium. They have limited computational and energy resource as they are usually small in size. In certain types of application, the information obtained from the WSNs has to be accurate and free from anomalies. They can be classified into data anomalies, network anomalies and node anomalies. To design an anomaly detection system for WSNs is a challenging task. In detection system usually detects anomalies such as viruses or network attack based on the signature in a centralised approach where the data are sent to a high performance server to be analyzed. Various anomaly detection techniques based on the three approaches described have been applied to WSNs to detect anomalies. However, most of these techniques are based on a normal model generated offline which does not change over time during detection. This is not sufficient for the time variant behaviour of the WSNs.

The rest of the paper is organized as follows: In section 2, a brief review of irregularity. The section 3, describes the proposed work and the section 4, detect anomalies system. The Results is given in section 5 and section 6 describes the Conclusion.

2. IRREGULARITIES IN WIRELESS SENSOR NETWORKS
In Wireless sensor network, irregularities can occur in the nodes, application data, transmission channels and networks and can be caused by the different types of errors and malicious attacks.

2.1 Data irregularity
Data irregularity in wireless sensor network can be caused by hardware faults or different types of errors and malicious attacks. Systematic errors are caused by changes in the operating conditions such as humidity and temperature, and calibration error of the sensor node. These errors can affect the sensor reading, resulting in inaccurate measurements leading to incorrect action taken.

2.2 Network irregularity
Network irregularity can be caused by different types of node failure or malicious attacks, unavailability of transmission radio. Most network irregularity found in the literature is attack-related as existing routing protocols can be easily exploited. A malicious node can attack a network by giving wrong routing information into the network or by sending an echo packet.

2.3 Signal irregularity
Physical communication medium can be greatly affected by noisy environments causing the signal to loss or attenuate. Radio irregularity and channel interference are common phenomena in wireless transmission that create transmission error in wireless sensor network. A collision can occur when two nodes simultaneously try to access the communication channel of the same frequency. A malicious node can listen

and learn the transmission cycle of wireless sensor network using TDMA and start a collision at each transmission cycle.

2.4 Node irregularity

Node irregularity is usually caused by physical attack or node failure. Sensor nodes are usually manufactured at low cost and can be easily damaged. They operate in an unattended location and are not tamper resistant making them easily to be under physical attacks. A node can be replicated using the cryptographic secret and Node identity extracted from the compromised or damage node. It allows the attacker to gain access to the network and corrupt data packet or even disconnect significant parts of the network. This kind of attack can be very difficult as the replicated node holding cryptographic secret can be considered as valid node.

3. PROPOSED WORK

An algorithm to increase detection sensitivity. It can detect the anomalies by applying various statistical distribution and distance functions to build the model based on the observed behaviour of the systems. It requires a certain measure criteria or threshold to identify the anomalies. This approach is very suitable for real world applications especially in WSNs where pre-labelled data are not easily available and the behaviour of the systems cannot be determined prior to deployment. In this paper the detection agent is installed in every node. It monitors the behaviour of neighbouring node within its transmission range locally to detect any abnormal behaviour. To perform real time anomaly detection, a decentralized detection approach has been proposed in using a rule based detection model installs in a monitor node. Monitor node listens promiscuously to neighbouring nodes within its transmission range to collect data necessary for anomaly detection.

4. DETECT ANOMALIES

In Algorithm, A's role is to decide whether Z_{k+1} is abnormal or not. Node A can overhear node B's transmission Z_{k+1} at time t_{k+1}. After estimating $\hat{x}_k^+$at time t_k, A can predict node B's transmitted value $\hat{x}_{k+1}^-$ at time t_{k+1} based on Equation (1). At time t_{k+1}, A overhears B's transmitted value z_{k+1} and compares $\hat{x}_{k+1}^-$ with Z_{k+1} to decide whether B is acting normally or not. If the difference between $\hat{x}_{k+1}^-$ and Z_{k+1} is larger than Δ, a predefined threshold, A then raises an alert on B. Otherwise, A thinks that B functions normal.

1. A time update-state equation is used to predict the state $\hat{x}_{k+1}^-$ at time t_{k+1}:

   $\hat{x}_{k+1}^- = F(\hat{x}_k^+)$ (1)

2. A measurement update-estimate update with measurement Z_{k+1}.This equation is used to update estimate with measurement Z_{k+1}:

   $\hat{x}_{k+1}^+ = \hat{x}_{k+1}^- + (Z_{k+1} - \hat{x}_{k+1}^-)$ (2)

Steps of Algorithm:

Assumption: Node A can overhear node B's transmission. A thinks that B is a normal node at and before time t_k

Input: Z_{k+1} transmitted by node B and overheard by node A.

Output: Whether A raises an alert on Z_{k+1}.

Procedure:

Step 1: At time t_k, A computes $\hat{x}_k^+$ based on Eq. (2) (note that $\hat{x}_k^-$ is stored in node A);

Step 2: A computes $\hat{x}_{k+1}^-$ based on x_k^+ using Eq.(1);

Step 3: A computes Diff $=|\hat{x}_{k+1}^- - Z_{k+1}|$;

Step 4: **if** (Δ < Diff) **then**

Step 5: A raises an alert on B;

Step 6: **else**

Step 7: A thinks that B functions normally;

Step 8: **end if**

5. SIMULATION AND RESULTS

5.1 Simulation Model

NS-2 simulator is used for performance evaluation. The network is a collection of 30 nodes deployed on square area of 800mx800m. Transmission range of each node is 250 m. For radio propagation model, a two-ray ground reflection model is used. In our simulations, we will use the RWP (Random waypoint) mobility model. Each node moves with a maximum speed randomly chosen from the interval 5 m/s and 15 m/s. Communication between nodes is modelled by CBR (Constant Bit Rate) traffic over UDP. A source generates packets of 512 bytes with a rate of five packets per second. A total of 20 connections were generated. They start at a time randomly chosen from the interval [0s, 100s] and still active until the end of simulation. We consider Random way point mobility (RWP) for mobility. We simulate the network at 30 nodes. In order to find the best mobility model we fix the CBR connection and pause time of each node.

5.2 RESULTS

The total number of nodes 30 in our simulation evaluation process. The simulation process every node is working in cooperation with each other to keep the network in communication.

The simulations are carried out for network densities of 30 nodes respectively. The area considered is 800m X 800m for stationary nodes and nodes with mobility of 10mps. Simulations are configured for the first node dead and all node dead estimations of both routing protocols with the metrics like battery capacity & energy consumed at the destination for stationary and nodes with mobility of 10mps respectively. Comparison of routing protocols constant bit rate (CBR) traffic patterns is used. The network contains variable CBR traffic connections and packet size of 512 bytes.

Table 1: Packets transmitted with and without Anomalies (Simulation time 15)

Simulation Time	Packets Without Anomalies	Packets With Anomalies
0	0	0
5	40	620
10	200	1490
15	440	1800

In figure 1, the numbers of nodes are 30 and simulation time is 15. When simulation time is 0 no packet is send with and without Anomalies but when simulation time is increasing a large difference between the packets transmission are seen with and without anomalies e.g. when the simulation time is 5 the number of transmitted packets without anomalies is 40

and the number of transmitted packets with anomalies is 620. When the less number of packets are transmitted the energy consumed is also less.

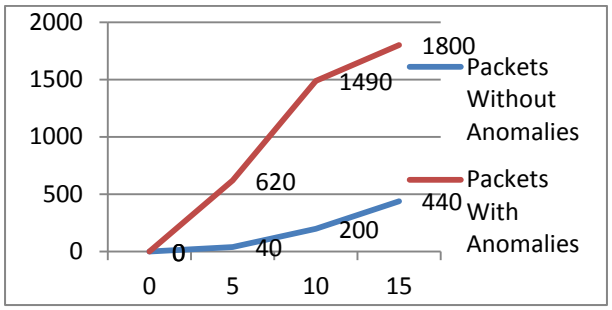

Figure 1 Simulation Time vs. Number of packets transmitted (with or without anomalies)

In figure 2, the numbers of nodes is 30 and simulation time is 15. When simulation time 0 no energy consumed with and without Anomalies but when simulation time is increasing the energy consumption is also increasing with respect to the simulation time e.g. when simulation time is 5 the energy consumed without anomalies is 23 and the energy consumed with anomalies is 122.

Table 2: Energy Consumed with and without Anomalies (Simulation time 15)

Simulation Time	Energy Consumed Without Anomalies	Energy Consumed With Anomalies
0	0	0
5	23	122
10	91	230
15	162	332

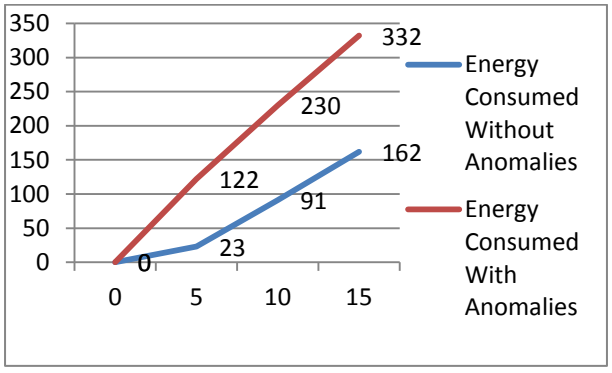

Figure 2 Simulation Time vs Energy Consumed (with or without anomalies)

In figure 3, the numbers of nodes are 30 and simulation time is 30. When simulation time is 0 no packet is send with and without Anomalies but when simulation time is increasing a large difference between the packets transmission are seen with and without anomalies e.g. when the simulation time is 5 the number of transmitted packets without anomalies is 50 and the number of transmitted packets with anomalies is 620. When the less number of packets are transmitted the energy consumed is also less.

Table 3: Packets transmitted with and without Anomalies (Simulation time 20)

Simulation Time	Packets Without Anomalies	Packets With Anomalies
0	0	0
5	50	620
10	200	1500
15	440	1840
20	643	2050

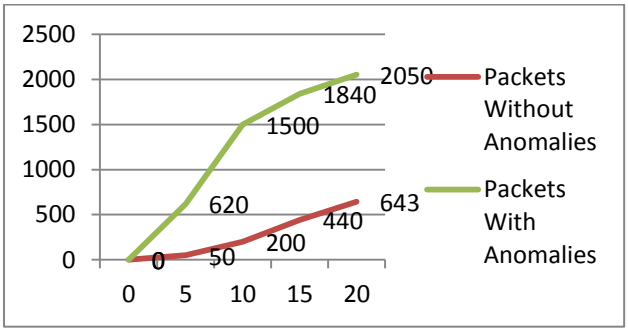

Figure 3 Simulation Time vs. Number of packets transmitted (with or without anomalies)

In figure 4, the numbers of nodes is 30 and simulation time is 15. When simulation time 0 no energy consumed with and without Anomalies but when simulation time is increasing the energy consumption is also increasing with respect to the simulation time e.g. when simulation time is 5 the energy consumed without anomalies is 24 and the energy consumed with anomalies is 125.

Table 4: Energy Consumed with and without Anomalies (Simulation time 20)

Simulation Time	Energy Consumed Without Anomalies	Energy Consumed With Anomalies
0	0	0
5	24	125
10	89	237
15	162	330
20	237	437

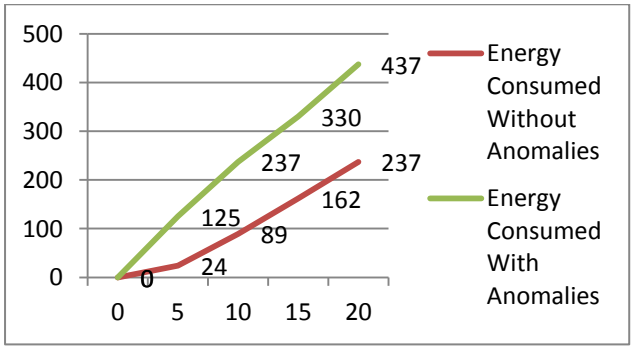

Figure 10 Simulation Time vs Energy Consumed (with or without anomalies)

Throughput

Data packet sent with Anomalies 375 and 347 were delivered by the throughput is 151.913 kbps but data packet sent without Anomalies 631 and 601 were delivered by the throughput. is 350.345 kbps. So we increase the network lifetime. We maximize the network lifetime and minimize the power consumption.

6. CONCLUSION AND FUTURE SCOPE

The primary aim of the thesis to described new efforts providing in wireless sensor networks. Many anomaly detection algorithms have been proposed that differ according to the information used for analysis, and the techniques applied to detect deviations from normal behaviour. Now developed an efficient algorithm to detection of sensor data irregularities in the wireless sensor networks. A sensor node monitors its neighbor's behaviour and establishes a normal Range of the neighbor's future aggregated values. An alert can be raised if the monitored value lies outside of the predicted normal range. In this work a scheme has been proposed to maximize the network lifetime and minimizes the power consumption during the detection of sensor data irregularities in the wireless sensor networks.

For future work we can improve the detection of sensor data irregularities in the wireless sensor networks. Developed more an efficient algorithm to detection of sensor data irregularities in the wireless sensor networks based on the better predicted normal range.

7. REFERENCES

[1] Yong Wang, Garhan Attebury and Byrav Ramamurthy "A Survey of security issues in Wireless Sensor Networks", *IEEE Communication Survey* 2006.

[2] V. Chandola, A. Banerjee, and V. Kumar. Anomaly Detection: A survey. ACM Computing Survey, 41(3):1{58, 2009.

[3] K. Ni, N. Ramanathan, M. Chehade, L. Balzano, S. Nair, S. Zahedi, E. Kohler, G. Pottie, M. Hansen, and M. Srivastava. Sensor network data fault types. ACM Transaction in Sensor Networks, 5(3):1{29, 2009.

[4] B. Parno, A. Perrig, and V. Gligor. Distributed detection of Node replication attacks in sensor networks. In Proceedings of the *IEEE Symposium on Security and Privacy*, pages 49{63, 2005}.

[5] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. Impact of radio irregularity on wireless sensor networks. *In Proceedings of the 2nd international conference on Mobile systems, applications, and services (MobiSys04*), pages 125{138, 2004.

[6] C. Karlof and D. Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. In Proceedings of the *First IEEE International Workshop on Sensor Network Protocols and Applications*. Pages 113 {127, 2003.

[7] Design and Implementation of Mobile Robot for Nodes Replacement in Wireless Sensor Networks* JANG-PING SHEU, KUN-YING HSIEH+ AND PO-WEN CHENG, 200

[8] H. Zhang and H. Shen, "Balancing energy consumption to maximize network lifetime in data-gathering sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 10, pp. 1526–1539, 2009.

[9] Y. Revah and M. Segal, "Improved algorithms for data gathering time in sensor networks II: Ring, Tree, and Grid topologies," *International Journal of Distributed Sensor Networks*, vol. 5, no. 5, pp. 463–479, 2009.

[10] K. Yuen, B. Liang, and B. Li, "A distributed framework for correlated data gathering in sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 578–593, 2008.

[11] S. Susca, F. Bullo, and S. Martinez, "Monitoring environmental boundaries with a robotic sensor network," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 2, pp. 288–296, 2008.

[12] M. Ma and Y. Yang, "Data gathering in wireless sensor networks with mobile collectors," in Proceedings of the *22nd IEEE International Parallel and Distributed Processing Symposium* (IPDPS '08), April 2008.

[13] A. Boukerche and X. Fei, "Adaptive data-gathering protocols with mobile collectors for vehicular ad-hoc and sensor networks," in Proceedings of the *4th IEEE International Conference on Wireless and Mobile Computing, Networking and Communication* (WiMob'08), pp. 7–12, 2008.