ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 3, March 2014

Extended Faulty Nodes and Event Identification for WSNs Using IDM and SMM

Anvy Francis p[1], Halice K Babu[2]

M-Tech, CSE, JCET, Ottapalam, India[1,2]

Abstract: Wireless sensor networks (WSN) have gained great popularity, mainly because they provide a low cost alternative to solving a great variety of real-world problem. Wireless Sensor Networks (WSN) are characterized by the dense deployment of sensor nodes that continuously observe physical phenomenon. In-network aggregation is an essential primitive for performing queries on sensor network data. Security in-network aggregation for wireless sensor networks (WSNs) is a necessary and challenging problem. The existing techniques involves identification of malicious activities when one neighbouring node is compromised. The proposed system introduce the integration of system monitoring modules and intrusion detection modules when multiple neighbouring faulty nodes occur in the network. And also propose an extended Kalman filter (EKF) based local detection mechanism to detect false injected data. Specifically, by collecting the information from its neighbours and by using EKF to predict their future states (actual in-network aggregated values).Each node aims at setting up a normal range of the neighbours' future transmitted aggregated values. This is challenging task because of potential high packet loss rate, time delay, harsh environment, and sensing uncertainty. Using aggregation function (average), obtain a theoretical threshold. By comparing this threshold value with the measured value, conclude the whether the event is malicious or not.

Keywords: Wireless sensor networks, IDS(Intrusion Detection System),EKF(Extended Kalman filter), In-Network Aggregation.

I. INTRODUCTION

A wireless sensor network (WSN**)** consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a main location. The more modern networks are bi-directional, also enabling control of sensor activity. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control, machine health monitoring, and so on. Applications of Wireless Sensor Networks (WSNs) range from military surveillance to environmental monitoring. WSNs applications, such as in health care, have recently also been envisioned to help our daily lives. However, sensor nodes in these applications could easily be compromised and can inject falsified data into the networks.

In-network aggregation is an important primitive which reduce the communication overhead and also save energy for WSNs. Many aggregation protocols have been introduced in earlier and their performance has been evaluated. However, only a few protocols such as encryption, authentication, and key management etc are secure in-network aggregation based on a prevention-based scheme. If a sensor node is compromised by an adversary, this adversary can take full control of the compromised node. It may inject falsified data readings or nonexistent readings into the WSN., so prevention-based techniques will be helpless in this situation. To overcome this problem, intrusion detection systems (IDSs), which serve as the second wall of protection, can effectively help to identify malicious activities. An Anomaly-Based Intrusion Detection System, is a system for detecting computer intrusions and misuse by monitoring system activity and classifying it as either normal or anomalous. The classification is based on heuristics or rules, rather than patterns or signatures, and attempts to detect any type of misuse that falls out of normal system operation.

To provide WSN security, the proposed system integrated both, System Monitoring Modules (SMM) and Intrusion Detection Modules (IDM), in each node. This integration helps to identify the classification between malicious events and important emergency events such as forest fire. For example, using IDM, when node A raises an alert on node B because of an event E, node A can further initiate investigation on event E with the help of SMM. Specifically, node A can wake up relevant sensor nodes around node B and request their opinions about event E. If the majority of sensor nodes think that event E could happen, node A can make a decision that event E is triggered by some emergency event. Otherwise, node A can suspect that event E is malicious.

Then design an Extended Kalman Filter (EKF) based local detection mechanism to detect false injected data. Specifically, by collecting the information of its neighbors about the event and using EKF to predict actual in-network aggregated values (states), For this consecutively observe the sensor nodes, and predict their future observed values based on previous values for each nodes.

If the measured value which is calculated during event occurs, lies below the threshold value which is predefined, the event will be emergency. Otherwise it will be malicious. If one of the neighbor has been compromised, it may not generate correct result [1]. This motivates our proposed local detection algorithms,

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 3, March 2014

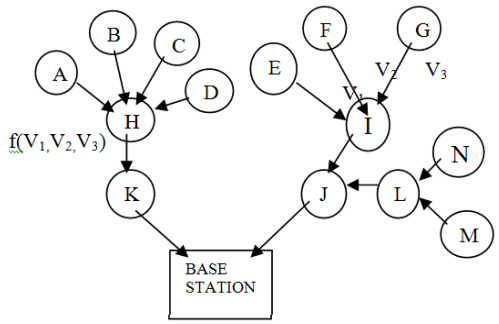

Fig.1.Aggregation tree

nodes close to each other can have spatially correlated observations, which can facilitate the collaboration of sensor nodes in proximity to differentiate between malicious events and important emergency events. This motivates us to integrate specialized SMM and IDM modules in order to achieve accurate detection results.

II. AGGREGATION MODEL AND ASSUMPTIONS

A. Aggregation Model

An aggregation tree can be modelled, to provide data aggregation[5]. Fig. (1) is one example of such an aggregation tree. In this nodes A, B, C, and D are sensor nodes, thus they obtain values and transmit these values to their parent node H. H aggregates (in the proposed paper take the average) the received values from nodes A, B, C , and D , and transmits the aggregated value further up to parent node K . The same operation is performed for nodes (E, F, G)→I→J and nodes(M, N)→L→J. These aggregation operations are performed based on the established parent-child relationship.. At last the base station collects all these data values and, if necessary, can transmit them across the Internet.

B. Assumptions

WSNs are mainly deployed to monitor emergency events. Here assume that the majority of nodes around some unusual events are not compromised. In anomaly based detection the behavior of the majority of nodes will be the normal system behavior. The number of such nodes is much larger than other nodes. When a sensor node is affected by an adversary, it will take all secret information about the sensor node. And the adversary can inject false data values even non-existent readings into the wireless sensor networks. The malicious data transmitted by a malicious node is significantly different from actual state. Thus the false data can disrupt the aggregation operation. The proposed systems do not assume time synchronization among nodes. It can tolerate the time inaccuracy caused by children nodes and parent nodes.

The promiscuous mode is supported by sensor nodes. By enabling promiscuous mode, in Fig (1) node F can overhear node I's transmissions. This facilitates the proposed neighbor monitoring mechanisms. For the purpose of saving energy, sensor node scheduling policies such as some sensor nodes goes to sleep mode during processing must be considered. But necessary sensor nodes could be wake up anytime when it will be required

III. NORMAL EVENT IDENTIFICATION USING EKF

The proposed system consist of two modules: Intrusion Detection Module (IDM) and System Monitoring Module (SMM). The functionality of the IDM is to detect whether monitored nodes are malicious nodes, while the functionality of the SMM is to monitor important emergency events. IDM and SMM need to be integrated with each other to work effectively. Each node contains very limited information. Since sensor nodes are prone to failure, it is very difficult to differentiate between emergency events sent by good nodes and malicious events. In this system to identify the original or malicious event the system have to collect the information from the surrounding nodes. For the IDM[2], node A collect the values from its neighbor's and compares this value with the normal range. If this value is greater than the normal range, either an event E happens or the neighbor N then becomes a suspect. To conclude, whether the node Nis a malicious node or event E is an important emergency event, node Ainitiates IDM and SMM module by waking up relevant sensor nodes around node N and requesting their opinions about event E. And if majority of sensor nodes tells that event E will be occur which is the malicious, the proposed system compares the values generated in two neighboring IDN-SMM module. If this value is approximately same or lies inside normal range it will be normal event otherwise it will be malicious event.

A. Extended Kalman Filter based Local Detection

1. Extended Kalman Filter:

The most general form of state space model is the non-linear model. The models are basically consist of two function 'F' and ' H' which govern the state propagation and measurements respectively. 'w'and 'v'are the process and measurement noises respectively, and k is the discrete time.

$x_{k+1}=F(x_k)+w_k(1)$

$z_k=H(x_k)+v_k=x_K+v_k(2)$

This is the actual model where as the linear state-space model is the model where the functions F and H are both linear in state and input. Based on a state-space model, Kalman Filter (KF) addresses a general problem of trying to estimate a state of a dynamic system perturbed by Gaussian white noise. But by using Extended Kalman Filter mechanism the system can set a proper process modeland measurement model for a specific WSN application and can utilize time update and measurement update equations to recursively process data. State indicates an actual value which is to be measured. State at a given instant of time is characterized by instant an eous values of an attribute of interest. The aggregation nodes will obtain values other than state values. Aggregation nodes or aggregators can only obtain measured values to estimate actual values. Aggregator calculate aggregated values consecutively. So the system can use a Discrete-Time Extended Kalman Filterin which a system state is estimated at a discrete set of $timest_k$, where k = 0, 1, .These discrete times correspond to the times at which a value is measured and a state is estimated. Actual aggregated values and a process model, given in (1),

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 3, March 2014

governs the evolution of this process. If WSNs are deployed to monitor the average indoor temperature of a building, $F(x_k)$ may be set tox_k. If the monitored temperature decreases gradually in a time period, one possible $F(x_k)$ may be set toδx_k, where δ is a positive value less than 1.

Measurement Modelis given in (2). $z_k$is the measured value at time t_k. For example, in Fig. 1, node I sends out an aggregated value $Z_k$at timet_k, node E, F, and G can overhear this value. $x_k \varepsilon R$ (R denotes the set of real numbers) is the state to be monitore date timet_k and represents the actual aggregated value of the area that aggregation node I covers. u_k follows a normal distribution with mean 0 and variance R, denoted as N(0,R), where R is the variance of u_k and this will be the measurement noise, representing noisy sensor measurements and various uncertainties in WSNs.

2. System equations:

Plot the actual, measurement, and estimate value using these system equations.

A Time Update - State Estimate Equation is used to predict the statex_{K+1}at time t_{K+1}.

$$x_{k+1}=F(x_k) \quad (3)$$

A Measurement Update - Kalman Gain Equation is used to Compute the Kalman Gain at time tk+1:

$$P_{K+1}=P_{K+1}(P_{K+1} + R_K)^{-1} =P_{K+1}/P_{K+1} + R_K \quad (4)$$

A Measurement Update - Estimate Update with Measurement Z_{K+1}Equation is used to update estimate with measurement Z_{K+1}

$$x_{K+1}=x_{K+1}+k_{K+1}(z_{k+1}-x_{k+1}) \quad (5)$$

For details about the derivation of these equations, refer to [3].

B. Threshold based Anomaly Detection Mechanisms:

A sensor node monitors its neighbor's behavior or values and establishes a normal range of the neighbor's future aggregated values. The creation of the normal range is centered on estimated values using EKF. An alert can be raised if the monitored value lies outside of the predicted normal range. This scheme is described in Algorithm. Here Δ is a predefined threshold. In Algorithm, A's role is to decide whether Zk+1 is abnormal or not. Node A can overhear node B's transmission Z_{k+1} at time tk+1. After estimating x_k^+at time tk, A can predict node B's transmitted value x_{k+1}at time t_{k+1}based on Equation (3). At time t_{k+1}, A overhears B's transmitted value $z_{k+1}$and compares x_{k+1} with z_{k+1} to decide whether B is acting normally or not. If the difference between x_{k+1} and z_{k+1}(denoted as Diff) is larger than Δ, aprede fined threshold, A then raises an alert on B. Otherwise, A thinks that B functions normally [4].

Algorithm: EKF based local detection algorithm
Input: z_{k+1} transmitted by node B and can be received by node A
Output: Whether A raises an alert on z_{k+1}
Procedure:

1. A computes posteriori estimate$(x_k^+)$of $x_k$based on (5) at time t_k;
2. A computes priori estimate(x_{k+1}) based on x_k^+ using (3);
3. A computes Diff= $\left| x_{k+1} - z_{k+1} \right|$;
4. If(Δ<Diff) then
5. A raises an alert on B;
6. Else
7. A thinks that B functions normally;
8. End if
9. Move to step 3 until all neighboring nodes are covered

Each sensor node transmits value $v_i$to its parent node based on a predefined aggregation protocol. Suppose that the expectation of each $v_i$is $E[v_i] = \mu_i$and the variance of each $v_i$is var $(v_i) = \sigma_i^2$

The proposed system is illustrated in figure2. Here nodes 11 and 13 generate false event and original event respectively. When a node 11 raises an alert on its parent node, it will collect information from neighboring nodes (8, 9, 10, 16, and 17) around node 11. And it will calculate estimate vale and compare with the threshold value. If it is an overhead value it will be suspicious. It may be malicious event. And the result sends to the base station. Similarly do the same procedure on node 13. If the neighboring nodes 8,9,10 are compromised, compare the neighboring IDM and SMM module. And send the alert to the base station.

IV. EXPERIMENTAL RESULTS

In this, the proposed system can use both live data and synthetic data to evaluate EKF based detection algorithms. Live data contain a limited number of situations which capture real world situations and whose parameters cannot be varied. Moreover, it may be difficult to obtain real attack data. In this situation, synthetic data, whose parameters under normal and abnormal situations can be carefully controlled, can offer advantages.

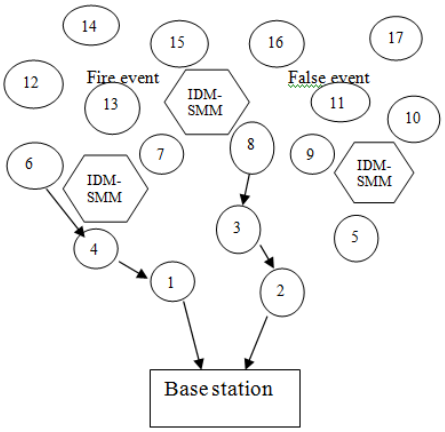

Fig 2.Collaboration between IDM and SMM

A. Simulation Results on Synthetic Data

Simulation Setup*:* By using network model , evaluate EKF based anomaly detection algorithms. Base on this plot the graph of the throughput and packet delivery ratio V_S time which is shown in fig 3 and 4.The green line shows situation of the proposed system. And the red line shows the situations of the existing system.. In Fig.(1), vi denotes

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 3, March 2014

the measured value by child node E,F,G, which is transmitted to node I for aggregation. Thus node I is an aggregator. Assume that the actual temperature value at nodes E,F,G is x_i. Since vi may be different from the actual value at node E,F,G. Then calculate the aggregate value of vi (i = 1, 2, ...n), denoted as z, and use z to estimate the actual aggregate value, x. Then set node I as a compromised node. That is, node I can inject falsified aggregated data into a network. Intuitively, it is easier to detect attackers that have larger variations from normal nodes. Therefore, introduce a concept degree of damage, denoted as D. D is defined as the difference between attack data and normal data. For example, in Fig. 1, the correct aggregated value by node I is z and the malicious aggregated value sent out by I is z_0, then we have $D = z_0 - z$. Then evaluate local detection schemes using different D based on this two metrics.

- False Positive Rate: It is measured over normal data items. Suppose that m normal data items are measured, and n of them are identified as abnormal. False Positive Rate is defined as n/m.

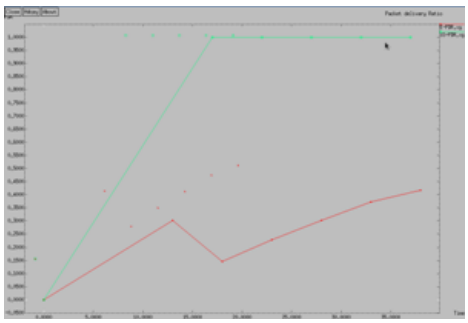

Fig 3. Packet Delivery ratio Vs Time

- Detection Rate: It is measured over abnormal data items. Suppose that m abnormal data items are measured, and n of them are detected. Detection rate is defined as n/m.

Under the same set of simulation parameters, obtain a certain amount of normal data items and a certain amount of malicious data items. For these data items, use different threshold values and measure the corresponding false positive rate and detection rate. In other words, for a given threshold value, by using algorithm, obtain one false positive rate and detection rate. Then apply different threshold values and obtain a set of false positive rates and detection rates.

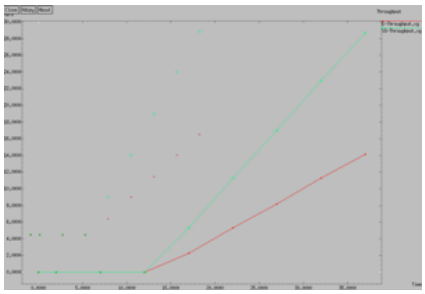

Fig 4.Throught Vs Time

And this graph indicates that the proposed system give high throughput and more packet delivery ratio than the existing system.

III. CONCLUSION AND FUTURE WORK

Security in wireless sensor networks is an important problem. To enhance WSN security in this system, first proposed that the integration or broadcasting of IDM and SMM to provide intrusion detection capabilities for WSNs. Then introduce EKF local detection mechanism to detect false injected data. Further demonstrated how the proposed IDM can work together with SMM to differentiate between malicious events and emergency events when multiple neighbouring nodes become malicious. Then evaluate the proposed schemes using synthetic data.Simulation results show that the proposed work is suitable to provide intrusion detection capabilities for secure in- network aggregation in wireless sensor networks. In the future, by considering more aspects or parameters of the proposed system, to make it more robust and effective.

REFERENCE

[1] Bo Sun,Xuemei Shan,Kui Wu, YangXiao, Senior Member IEEE, Nenghai Yu, and FenghuaLi"Anomaly detection based secure in-networkaggregation for wireless sensor networks", ieeesystems journel, Vol. 7, No. 1, March 2013

[2] S. Marti, T.J. Giuli, K. Lai, and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks,"ACM Mobicom'00, Boston, MA, Aug. 2000, pp. 255-265.

[3] M. S. Grewal, A. P. Andrews, "Kalman Filtering:Theory and Practice Using MATLAB," Wiley, Jan. 2001, ISBN: 0-471-39254-5.

[4] D.C. Montgomery, "Introduction to Statistical Quality Control," John Wiley & Sons, ISBN: 0-471-65631-3.

[5] G. Wang, W. Zhang, G. Cao, and T. La Porta, "OnSupporting Distributed Collaboration In SensorNetworks," IEEE Milcom 2003, Boston,MA, 2003

BIOGRAPHIES

Anvy Francis P received her BTECH Degree in Computer Science and Engineering from IES college of engineering Chittilappilly, Thrissur, which is affiliated to Calicut university Thenhipalam, Kerala. Now currently pursuing her Final Year Master's of Engineering in Computer Science and Engineering from Jawaharlal College of Engineering and Technology, which is affiliated to University of Calicut, Kerala. Her area of Interest is Digital Image processing, Computer Networks, Compiler Design.

Halice K Babu received her BTECH degree in IT under University of Calicut, Kerala and M.E Degree in Embedded System from Anna University, Chennai, Tamil Nadu. She is currently working as an Assistant Professor in the Department of Computer Science and Engineering at Jawaharlal College of Engineering and Technology, which is affiliated to university of Calicut. Her area of Interest is Digital Image processing, Embedded System, Data Mining, Computer Networks.