

Efficient Filtering and Location Detection Against Internal Threats in WSN

Jeena Elezabeth Cheriyan Dept. of CSE P.S.N.A College of Engineering & Technology, Dindigul, India.

S. Sathees Babu Dept. of CSE P.S.N.A College of Engineering & Technology, Dindigul, India. K. Balasubadra Dept. of IT RMD College of Engineering & Technology, Chennai, India

Abstract— Improving the security in wireless sensor networks (WSNs) is an important and tedious task. Node compromise is a major threat which allows an internal attacker to corrupt the network by injecting false data. This work illustrates how detection of false nodes using Extended Kalman Filter(EKF) mechanism work together with prevention against internal attackers using Timing control method. With the help of EKF algorithm to filter out false data values, the sensed value is compared against a predefined threshold in the nodes. A combination of CUSUM GLR algorithm is used for effective detection sensitivity of malicious nodes. In addition to previous approaches to mitigate internal threats, a timing control approach is also presented. In this method, sink node opens only for a specific time slot and then goes to sleeping state. This is done by calculating time slot at which high signal noise ratio is monitored. At this time period, it avoids all incoming signals. Thus it can prevent the network from false signals against internally malicious nodes.

Keywords - System monitoring modules (SMM), Extended Kalman Filter (EKF), cumulative summation(CUSUM), Generalized likelihood ratio (GLR).

I. INTRODUCTION

Wireless sensor networks (WSNs) have gained attention recently due to their huge significance in both military and civilian operations. They are utilized in hostile and unattended environments such as battlefield, weather monitoring operations. A typical WSN consists of a base station, sink node and surrounding nodes. The child nodes sense the environment and send values to the sink which aggregate and send the final report to the base station. Also, since WSNs have the broadcast nature, an adversary can easily overhear, known as a passive attacker, and active intruder will inject false messages. When WSN is utilized in open and possibly averse environments, intruders can easily launch denial of service(DoS) attacks. This could damage the sensors or avail sensitive data such as encrypted keys, address of secret data etc. Thus, insider threats poses serious threat to wireless network by which an intruder can access the private data.

In this work, we demonstrate how to locate and detect the malicious nodes which sends false signals to the sink. First, we propose the Extended Kalman Filter algorithm [1]. The false injected data is detected using EKF (Extended Kalman Filter) mechanism. This is done using System monitoring modules (SMM) by tracking the behavior of its nearby nodes which gets monitored and next possible state is predicted using EKF. It is difficult to find out which are emergency events and faulty events. A series of neighbor nodes' expected values in the succeeding states is being stored by each node. Then analysis is done to calculate the threshold value under different functions for aggregation using average of aggregation, sum, max, and min. An overheard value and normal range is compared using a threshold mechanism to decide whether there is a notable difference A sensor node when compromised by an adversary, this adversary can take full control of the compromised node. It then disrupts sensed data by changing the data readings in the WSN. It is assumed that data that is falsified is transmitted by a faulty node is very different from the state that is, the actual value so that falsified data can effectively other nodes also malicious. Environment noise, inaccuracy in sensing, time difference between children and parent nodes make this task tedious.

Second, a combined algorithm with cumulative summation (CUSUM) and generalized likelihood ratio (GLR) is used to perform genuine location detection in WSNs where cumulative sum of the deviations between sensed values from environment and estimated value is utilized. This when used along with filtering algorithm improves the throughput. Finally, a timing control algorithm [3] is introduced along with above mentioned algorithms here to prevent false signals to the sink node. In this method, sink node opens only for a specific time slot otherwise it will be in sleeping state. This is done by calculating time slot at which high signal noise ratio is monitored. At this time period, it avoids all incoming signals including the malicious signals which when injected could corrupt the network. Thus it can prevent the wireless sensor network from false signals within the sleeping time period.

II. RELATED WORK

The problem of internal threats was addressed by Zhang and Jajodia[2]. An attacker can gain control over the network, when a node is injected with false data and can impose false activities. These attacks send wrong information in the

network or even alter network and leads to false alarms. Karlof and Wagner have demonstrated network layer attacks and detailed about corrupted information in routing and denial of service attacks, replication of nodes, black grey sink holes. The work by Wagner used statistical estimation attack in which a mathematical framework is presented to evaluate security of different aggregation algorithm for more resilient aggregation schemes against false data injection. Nodes which lie as intermediate ones just forward input packets such as in normal routing. But these nodes can damage incoming packets by coding and final coded packets are forwarded which also makes it malicious. In pollution attack, attackers inject false packets into the network. Thus forwarded packets get falsified and this leads to malicious data forwarding in the network which will pose a serious security threat.

Jajodia, Setia et al [5] proposed scheme that enable the base station to verify the authenticity of a sensing report that it has received as long as the number of compromised sensor nodes does not exceed a certain threshold. Further, this attempts to filter out false data packets injected into the network by compromised nodes before they reach the base station, thus saving the energy for relaying them. Zhang and Liu[6] presented the idea of location-based keys in which private keys of single nodes is binded to both their IDs and physical locations. Bayesian algorithm uses statistical approach to deal with the possibility of measurement faults in sensors. These efforts have become a great support for target finding and fault data detection. Detection using kalman filter and cumulative summation mechanisms have also been widely used in many applications. For example, in the context of WSNs, KF was used to enable accurate target tracking [7]. Unlike above techniques, proposed scheme aims at addressing internal threats using effective algorithms mentioned above. KF and CUSUM have not yet been applied to secure WSN aggregation services. This paper relies on succeeding values of node in states with the help of nearby nodes and can cooperate with existing mechanisms to prevent attacks in network.

III. EXTENDED KALMAN FILTER DETECTION

A sensor node establishes a normal range of the neighbor's future aggregated values by monitoring neighbor's behavior. Normal range is calculated based on values using EKF. If the sensed value lies outside of the predicted normal range, then an alert is raised. In the 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 based \hat{x}_{k+1} at time t_{k+1} based on (3). 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. Else, A thinks that B functions normally.

Now, at time t_k , to predict the actual value x_{k+1} , a node needs two values:

1) A priori estimate \hat{x}_{k+1}^- which can be obtained based on (3).

2) The measured value z_{k+1} that can be overheard. EKF can provide a accurate prediction of neighbours' future values. Now we present EKF based location detection algorithm.

Algorithm 1 EKF based local detection algorithm

Assumption Node X can overhear node Y's transmission. X thinks that Y 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} 1: At time t_k , A computes \hat{x}_k^+ (\hat{x}_k^- is stored in node A); 2. A computes \hat{x}_{k+1}^- based on \hat{x}_k^+ using (3); 3. A computes Diff = $|\hat{x}_{k+1} - z_{k+1}|$; 4. if (Δ < Diff) then 5. A raises an alert on B; 6. else 7. A thinks that B functions normally; 8.end if

In the state space model, actual aggregated values form a dynamic process, and a process model given by,

 $x_{k+1} = f(x_k) + w_k$ (1) where x_k represents the actual value at time t_k . F is a function relating x_{k+1} to x_k and w_k is the process noise at time t_{ν} .

Measurement model is given by,

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

where zk is the measured value at time tk. H is the function relating xk to zk the function relating x_k to z_k and v_k is the measurement noise at time t_k . System equations is given by,

> $\widehat{x}_{k+1}^{-} = \mathbf{F}(\widehat{x}_{k}^{+})$ (3)

IV. CUSUM GLR BASED LOCAL DETECTION

An EKF based approach at times neglects the information given by the entire sequence of measured values. For example in Algorithm 1 if an attacker continuously injects z_{k+1} with small deviations, this leads to a small Diff. A relatively large Δ can make an EKF based approach insensitive to these kinds of attacks because this approach only uses information available at a previous time instant. An algorithm combining CUSUM and GLR is used which utilizes the cumulative sum of deviations between measured values and estimated values. The algorithm is based on following parameters.

Algorithm 2 CUSUM GLR based local detection algorithm

Assumption Node A can overhear node B's transmission. A thinks that B is a normal node at and before time tk **Input** A sequence of z_{k+1} transmitted by node B and overheard by node A **Output** Whether node A raises an alert on z_k 1: Compute $y_k = z_k - \hat{x}_k^-$ at time t_k . 2. Compute $\hat{\mu}_1 = \frac{1}{w} \sum_{i=k-w+1}^k y_i$ when $k \ge w - 1$ 3. $S_N = \frac{b}{\sigma} \sum_{i=0}^N \left(y_i - \mu_0 - \frac{v}{2} \right) = \frac{b}{\sigma} \sum_{i=0}^N \left(y_i - \frac{v}{2} \right)$ 4. if $(S_N > h)$ then 5. A raises alert on B; 6. else 7. A thinks that B functions normally; 8. end if

Consider a sequence of observed random variables y0, y1..., y_k with a probability density $p_{\theta}(y)$ depending on only one scalar parameter θ .

The log-likelihood ratio is defined by,

 $s_k = \ln \frac{p_{\theta 1}(y_k)}{p_{\theta 0}(y_k)}$ (4) s_k shifts from a negative value to a positive one when a change occurs in parameter θ

 $S_N = \sum_{i=0}^N s_i$ (5) This S_N incorporates the cumulative sum of s_k , S_N can be used to detect the change in y_k .

V. TIMING CONTROL ALGORITHM

There are three different nodes, target node, sensor node and sink node. The information received at the sink node over the wireless channel can be further analyzed by a control server or a human operator. Based on the content of the information, the sink node may have to send commands or queries to the sensor nodes. Therefore, the sink node only opens at a special time period other time is in sleeping state and ignores any coming signals such that it can protect the network from the internal attacks within the sleeping time period.

Algorithm 3 Timing Control Algorithm

Input p^{req}d (detection requirement), K (number of available sensors) r_d (detection radius), d (t, s) the distance between the target's position and **B**. Output u (deployment vector) and highest signal noise time and location. 3. For k = K - 1:-1:0 do 4. Evaluate G_k 5. end for 6. Initialization: k = 0, u = 07. while $k \leq K$ do 8. Find set of grid points with unsatisfied detection requirements, (i: $p_k d(i) \ge p^{req} d(i)$ 9. Set $x_{k-1}(i) = 0$ 10. Calculate the control vector $\mathbf{u}\mathbf{k} = -\mathbf{G}_{\mathbf{k}}\mathbf{x}_{\mathbf{k}}$ 11. Find the S/N and index jmax, where $j_{max} =$ max_{index} (uk). 12. Update the deployment vector (i.e. $u(j_{max}) = 1$) 13. Calculate $m_k = Bu$ 14. Calculate time t_k 15. Increment number of sensors in the grid, k = k + 116. End while

VI. EXPERIMENTAL EVALUATION

EKF and CUSUM GLR algorithms are implemented using network simulator, ns2. The purpose of the simulation is to filter out the malicious node and to find out whether an alert raised is genuine or not. The node size is set to 100 nodes. This to monitor a small geographic area for temperature values. Each node of the network is equipped with a 2 Mbps 802.11 radio with an omnidirectional antenna. Nodes will have 250m as the transmission range and 550m as their sensing range. The two-ray radio propagation model is used. The interference queue length is chosen as 50 packets in each node. We take the packet size as 1024 bytes at a packet rate of 8 packets per second. The minimum speed is taken as 5 m/s whereas the maximum speed is 8m/s. We collect data from the simulation run of 100 seconds.

In EKF algorithm, the parameters used are diff, a predefined threshold value is calculated using sum of aggregations or min, max value of aggregation. v_k is the measurement noise and w_k is the process noise at time t_k , priori and posteriori values already set in the node. The parameters used in CUSUM GLR algorithm are s_k , the ratio of observed values, s_N , the cumulative sum of deviations. h is the predefined threshold value, y_k is the observed value at t_k and $\hat{\mu}_1$ is the mean of observed values. This is used for comparing observed value with threshold value. The performance of both the algorithms can be evaluated by comparing the throughput. The graph shown below is plotted with throughput in the y-axis and number of nodes in x-axis. As the number of nodes increases, effective filtering of malicious nodes can be made possible.



Fig:1 Performance of EKF and CUSUM



Fig:2 Performance of Timing Control Algorithm

VII. CONCLUSION

Network security plays a significant role in technological advancement. The proposed model has a potential to be used as an effective and strong solution against the internal attacks. This work presents a combined effort of location detection and prevention of false nodes by using filtering and timing control algorithms. First, Extended Kalman Filter(EKF) along with CUSUM GLR algorithm is used to increase the detection sensitivity even if the adversary injects small amount of data into the network. In addition to the above methods, a timing control algorithm is also presented for effective prevention of false signals to the sink node. The sink node will be in sleeping state in high signal noise ratio conditions and avoids all incoming signals at this time period. This demonstrates the overall capability of the work to act against insider attacks in wireless sensor network.

ACKNOWLEDGEMENT

We would like to express our gratitude to all those who gave us the possibility to complete this paper.

REFERENCES

- [1] Bo Sun, Member, IEEE, Xuemei Shan, Kui Wu, SENIOR Member, IEEE, and Yang Xiao, Senior Member, IEEE, "Anomaly Detection Based Secure In-Network Aggregation for Wireless Sensor Networks" in IEEE SYSTEMS JOURNAL, VOL. 7, NO. 1, MARCH 2013.
- [2] Lei Zhang, Honggang Zhang, Mauro Conti, Roberto Di Pietro, Sushil Jajodia, Luigi Vincenzo Mancini "Preserving privacy against external and internal threats in WSN data aggregation", in Telecommun Syst (2013) 52:2163–2176 DOI 10.1007/s11235-011-9539-8.
- [3] Annapoorna Rao, Priyanka Singh, Shruthi R and Syeda S. Rubbani, "Defending mechanism to securenodes from internal attack in wsn" in International Conference on Advances in Computer and Electrical Engineering Nov. 17-18, 2012 Manila (Philippines).
- [4] Ahmad Ababnah, Balasubramaniam Naatarajan, "Optimal control based strategy for sensor deployment." IEEE Tran. On Systems, Man, and cybernetics, Part A: Systems and Humans, vol. 41, no. 1 Jan. 2011.
- [5] S. Zhu, S. Setia, S. Jajodia, and P. Ning, "An interleaved hop-by hop authentication scheme for filtering false data injection in sensor networks" in Proc. IEEE Symp. Security Privacy, pp. 260–272, May 2004.
- [6] Y. Zhang, W. Liu, W. Lou, Y. Fang, "Location-based compromise tolerant security mechanisms for wireless sensor networks," IEEE Journal on Selected Areas in Communications, Special Issue on Security in Wireless Ad Hoc Networks, Vol. 24, No. 2, pp. 247-260, February 2006.
- [7] J. Lin, L. Xie, and W. Xiao, "Target tracking in wireless sensor networks using compressed KF," Int. J. Sensor Netw., vol. 6, nos. 3–4, Nov. 2009
- [8] J. Staddon, D. Balfanz, and G. Durfee, "Efficient tracing of failed nodes in sensor networks," in WSNA 2002, pp. 122-130, Atlanta, USA.

- [9] Ochirkhand Erdene Ochir, Marine Minier, Fabrice Valois, and Apostolos Kountouris, "Resiliency of Wireless Sensor Networks: Definitions and Analyses", 2010 17th International Conference on Telecommunications, pp828-835.
- [10] G. Werner-Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, and M. Welsh, "Deploying a wireless sensor network on an active volcano," IEEE Internet Computing, vol. 10, pp. 18-25, 2006.
- [11] Hung-Min Sun, Chien-Ming Che, and Ying-Chu Hsiao, "An efficient countermeasure to the selective forwarding attack in wireless sensor network," IEEE Region 10 Conference (TENCON), 2007, pp.1-4.
- [12] Fang Liu, Xiuzhen Cheng, and Dechang Chen, "Insider Attacker Detection in Wireless Sensor Networks," IEEE International Conf. on Computer Communications (INFOCOM), May 2007, pp. 1937-1945.
- [13] Jing Dong, Reza Curtmola, and Cristina Nita Rotaru, "Practical Defenses Against Pollution Attacks in Intra-Flow Network Coding for Wireless Mesh Networks," WiSec'09, March 16-18, 2009, zurich, Switzerland, pp111-122.