International Journal of Recent Engineering Science (IJRES),
ISSN: 2349-7157, *volume1 Issue7 July2014*

Anomaly Detection Based Secure in – Network Aggregation for Wireless Sensor Networks

K.Rajesh
ECE, Asst. Professor, RVS College of Engineering and Technology, India

Abstract— *Secure in-network aggregation in wireless sensor networks (WSNs) is a necessary and challenging task. This project first proposes integration of system monitoring modules and intrusion detection modules in the context of WSNs. And propose an extended Kalman filter (EKF) based mechanism to detect false injected data. Specifically, by monitoring behaviors of its neighbors and using EKF to predict their future states (actual in-network aggregated values), each node aims at setting up a normal range of the neighbors' future transmitted aggregated values. This task is challenging because of potential high packet loss rate, harsh environment, and sensing uncertainty. The project illustrates how to use EKF to address this challenge to create effective local detection mechanisms. Using different aggregation functions (average, sum, max, and min), presents how to obtain a theoretical threshold. The project further applies an algorithm of combining cumulative summation and generalized likelihood ratio to increase detection sensitivity. To overcome the limitations of local detection mechanisms, it illustrates how The proposed local detection approaches work together with the system monitoring module to differentiate between malicious events and emergency events. The project simulates to evaluate local detection mechanisms under different aggregation functions.*

Keywords – *Channel resolution, Garbage output, constant input, power consumption, omni-directional antenna.*

I. INTRODUCTION

Area monitoring is a common application of WSNs. In area monitoring, the WSN is deployed over a region where some phenomenon is to be monitored. A military example is the use of sensors to detect enemy intrusion; a civilian example is the geo-fencing of gas or oil pipelines. When the sensors detect the event being monitored the event is reported to one of the base stations, which then takes appropriate action. Similarly, wireless sensor networks can use a range of sensors to detect the presence of vehicles ranging from motorcycles to train cars. To protect humans and the environment from damage by air pollution, it is of the utmost importance to measure the levels of pollutants in the air. Real time monitoring of dangerous gases is particularly interesting in hazardous areas, as the conditions can change dramatically very quickly, with serious consequences.

The measurement of gas levels at hazardous environments requires the use of robust and trustworthy

equipment that meets industrial regulations. Outdoor monitoring of air quality requires the use not only of accurate sensors, but also rain & wind resistant housing, as well as the use of energy harvesting techniques that ensure extended autonomy to equipment which will most probably have difficult access The term Environmental Sensor Networks, has evolved to cover many applications of WSNs to earth science research. This includes sensing volcanoes, oceans, glaciers forests, etc. Some of the major areas are listed below.

II. PROPOSED SYSTEM

In this project, we first propose integration of system monitoring modules and intrusion detection modules in the context of WSNs. We propose an extended Kalman filter (EKF) based mechanism to detect false injected data. Specifically, by monitoring behaviors of its neighbors and using EKF to predict their future states (actual in-network aggregated values), each node aims at setting up a normal range of the neighbors' future transmitted aggregated values. This task is challenging because of potential high packet loss rate, harsh environment, and sensing uncertainty. We illustrate how to use EKF to address this challenge to create effective local detection mechanisms. Using different aggregation functions (average, sum, max, and min), we present how to obtain a theoretical threshold. We further apply an algorithm of combining cumulative summation and generalized likelihood ratio to increase detection sensitivity. To overcome the limitations of local detection mechanisms, we illustrate how our proposed local detection approaches work together with the system monitoring module to differentiate between malicious events and emergency events. We conduct experiments and simulations to evaluate local detection mechanisms under different aggregation functions.

There are many research efforts that address aggregation problems in WSNs. However, none of the aforementioned protocols considers secure aggregation problems until recently. Hu and Evans tackled the problem of information aggregation in which one node is compromised. Their protocol might be vulnerable if both a child node and its parent node are compromised. Yang et al. proposed a secure hop-by-hop data aggregation protocol based on principles of divide-and conquer and commit-and-attest. Przydatek et al. proposed an aggregate-commit-prove framework to design secure data aggregation protocols. Chan et al. presented an optimally secure aggregation scheme for arbitrary aggregator topologies and multiple malicious nodes. Wagner used statistical estimation to design more resilient aggregation

schemes against malicious data injection attacks. In his work, a mathematical framework is presented to formally evaluate security of different aggregation algorithms. However, no detailed simulations and experiments are carried out in . Moreover, does not consider in-network aggregation. Our work improves over in these aspects. Wu et al. proposed a secure aggregation tree to detect and prevent cheating in WSNs, in which the detection of cheating is based on topological constraints in a constructed aggregation tree. There are some resilient aggregation algorithms aiming to increase the likelihood of accurate results when WSNs are prone to message loss and node failure . Also a number of proposed protocols aim to ensure the secrecy and authentication of data in WSNs. Several protocols are proposed to filter false data in WSNs. Generally, they utilize different key distribution mechanisms to develop filtering capabilities. In these research efforts, different sensing reports are validated by message We assume that promiscuous mode is supported by sensor nodes. By enabling promiscuous mode, when one node F is within the radio transmission range of another node, node *F* can overhear node I is transmissions. This facilitates our proposed neighbor monitoring mechanisms. For the purpose of saving node energy, there have been extensive research efforts on various kinds of sensor node scheduling policies, in which a minimum number of nodes remain awake to satisfy a certain degree of coverage. Therefore, we assume that sensor nodes may go to sleep. However, we also assume that necessary sensor nodes could be woken up anytime once required. We realize that in a real system, it may need nonzero time to allow a node to become fully functional. Therefore, our proposed scheme may not work very well if the period of the attackers' false injection is very short. This can be an open problem, and will be explored in our or our future work.

Consecutive observations of sensor nodes are usually highly correlated in time domains . This correlation, along with the collaborative nature of WSNs, makes it possible to predict future observed values based on previous values. This motivates our proposed local detection algorithms. Furthermore, since WSNs are usually densely deployed, 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 SMM and IDM in order to achieve accurate detection results.

To utilize data aggregation, an aggregation tree is often built first. Fig.3.3.2.1 is one example of such an aggregation tree. In Fig.3.3.2.1 , A,B,C, and D perform sensing tasks, obtain valuesand transmit them to their parent node *H*. *H* aggregates (min, max, sum, average, etc.) the received values from A,B,C, and D, and transmits the aggregated value further up to node

K. The same is true for operation (E, F,G) → I → J and operation (M,N) → L → J. These aggregation operations are performed based on the established parent–child relationship, which can be modeled using Fig.3.3.2.2. In Fig.3.3.2.1, the base station collects all these data and, if necessary, can transmit them across the Internet.

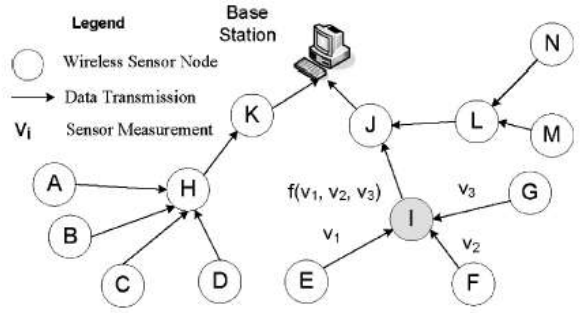

Fig.1: Example aggregation tree.

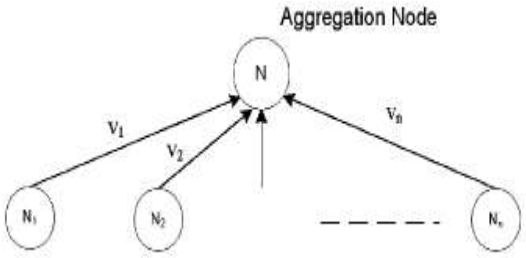

Fig.2: Aggregation model

WSNs are often deployed to monitor emergency events such as forest fires. We do not assume time synchronization among nodes. Our proposed approach can tolerate the time inaccuracy caused by child nodes and parent nodes. In the context of WSNs, time synchronization still incurs expensive operations We assume that promiscuous mode is supported by sensor nodes. By enabling promiscuous mode, when one node, e.g., F

in Fig. 3.3.2.1 is within the radio transmission range of another node, e.g., I, node *F* can overhear node *I*'s transmissions. This facilitates our proposed neighbor monitoring mechanisms. For the purpose of saving node energy, there have been extensive research efforts on various kinds of sensor node scheduling policies, in which a minimum number of nodes remain awake to satisfy a certain degree of coverage. Therefore, we assume that sensor nodes may go to sleep. However, we also assume that necessary sensor nodes could be woken up anytime once required. We realize that in a real system, it may need nonzero time to allow a node to become fully functional. Therefore, our proposed scheme may not work very well if the period of the attackers' false injection is very short. This can be an open problem, and will be explored in our future work. In this project, to simplify the study, and similar to many scheduling papers in sensor networks, we assume that any node can be woken up anytime immediately.

However, in a real system in practice, it may need nonzero time to allow node to become fully functional so that the proposed waking up- other node scheme may not work very well if the period of the attackers' false injection is very short. Such a problem can be an open topic that needs further study. Please note that if one node is in sleep mode, this node does not need to be in promiscuous mode. Existing research in self-monitoring for sensor networks can be integrated with our solution so that each active communication link can be monitored by nodes in the WSN. Moreover, in order to monitor sensor behaviors, there is an inevitable tradeoff of adopting promiscuous mode. Actually, any related work in this aspect cannot avoid this.

When 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. We also assume that falsified data transmitted by a compromised node is significantly different from the state (the actual value, for example, the actual monitored average temperature) so that falsified data can effectively disrupt aggregation operations. If the adversary only injects a limited number of falsified data that are slightly different from true aggregated values, this will not cause significant impact on deployed applications. Therefore, we will also consider an attack model that an adversary continuously forges falsified data with small deviations. We assume that the majority of nodes around some unusual events are not compromised. It will become an open research problem if this assumption does not hold.

Our proposed protocol is equipped with two modules: IDM and SMM. The functionality of the IDM is to detect whether monitored nodes are malicious insider nodes, while the function ality of the SMM is to monitor important emergency events. Note that SMM is a necessary component for most of WSN applications. IDM and SMM need to be integrated with each other to work effectively. Relying on local detection alone is not desirable because each node has only very limited information available. Furthermore, since sensor nodes are prone to failure, it is very difficult to differentiate between emergency events sent by good nodes and malicious events. In our proposed scheme, whenever IDM and SMM detect some abnormal events, they need to request the collaboration of more sensor nodes around the events to make a final decision. For the IDM, our general idea is like the mechanism proposed in. Node *A* promiscuously overhears its neighbor's transmitted aggregated value and compares it with the predicted normal range. If the overheard value lies outside the normal range, either an event *E* happens or the neighbor *N* then becomes a suspect. To tell whether node N is a malicious node or *E* is an important emergency event like the breakout of a forest fire, A initiates the collaboration between IDM and SMM by waking up relevant sensor nodes around N and requesting their opinions about E. Please note that our proposed detection solution and the solution adopted in are completely different.

Many challenges exist when we try to predict the normal range of in-network aggregated values in a lightweight manner. First, it is difficult to achieve actual aggregate values because of many sources of potential uncertainties. WSNs suffer from a high packet loss rate. For example, based on , in an in-building environment, with 62 motes deployed with the granularity of one mote per office, at a low load of 0.5 packet per second, there is around 35% of links whose packet loss is worse than 50% at a medium access control layer. There fore, even a reasonable link layer loss recovery is unable to mask high packet losses. For aggregation protocols, the lack of time synchronization among children and parent nodes may make aggregation nodes use different sets of values for aggregation. The complexity of existing aggregation protocols also contributes to the challenges of modeling in-network aggregated values. In it shows that for periodic aggregation, timing, i.e., how long a node waits to receive data from its children (downstream nodes in respect to the information sink) before forwarding data onto the next hop plays a crucial role in the performance of aggregation algorithms in the context of periodic data generation. Furthermore, individual sensor readings are subject to environmental noise. To demonstrate this, we set up a simple one-hop WSN test bed, in which node A periodically transmits sensed values to a base station. Node A consists of a MICA2 mote and a MTS310 sensor board . In a lab setting, we measure the collected data. We conduct a further experiment to demonstrate the uncertainty of the average aggregation function. In this experiment, we deploy four sensors to send their sensed temperature to an aggregation node B. B periodically computes the average of the received values. and (b) illustrates that data captured from a physical world and the aggregated values based on these data tend to be noisy. Sensor nodes suffer from stringent resources, which prevent the usage of some powerful yet expensive estimation and prediction approaches. To enable neighbor monitoring mechanisms, we need a lightweight scheme that can be efficiently executed by sensor nodes. In this respect, we use an approach based on EKF for each node to predict and estimate Future values of its neighbors, as we detail in the next section. to simplify the study, and similar to many scheduling papers in sensor networks, we assume that any node can be woken up anytime immediately. However, in a real system in practice, it may need nonzero time to allow node to become fully functional so that the proposed waking up- other node scheme may not work very well if the period of the attackers' false injection is very short. Such a problem can be an open topic that needs further study. Please note that if one node is in sleep mode, this node does not need to be in promiscuous mode. Existing research in self-monitoring for sensor networks can be integrated with our solution so that each active communication link can be monitored by nodes in the WSN. Moreover, in order to monitor sensor behaviors, there is an

inevitable tradeoff of adopting promiscuous mode. Actually, any related work in this aspect cannot avoid this.

Local detection alone is not enough. WSNs are often deployed to monitor emergency phenomena (like the breakout of a forest fire), about which good nodes can trigger important events and generate unusual yet important information. Also, the error prone nature of sensor nodes may make even normal sensor nodes faulty and generate abnormal information. Therefore, local detection alone suffers from a high false positive rate. Node collaboration is necessary for sensor networks to make correct decisions about abnormal events. Therefore, for WSNs, IDM and SMM need to integrate with each other to work effectively. When node A raises an alert on node *B* because of some event E, to decide whether E is malicious or emergent, A may initiate a further investigation on E by collaborating with existing SMMs. WSNs are usually densely deployed to collaboratively monitor some events. To save energy, some sensor nodes are periodically scheduled to sleep. Based on this, node A can wake up those sensor nodes (denoted as co detectors in around B and request from these nodes their opinions on the behavior of E. Because the majority of sensor nodes around the investigated event E are not compromised, after A collects the information from these nodes, if A finds that the majority of sensor nodes think that event E may happen, A then makes a decision that E is triggered by some emergency events. On the other hand, if A finds that the majority of sensor nodes think that event *E* should not happen, A then thinks that E is triggered by either a malicious node or a faulty yet good node. In this way, A can continue to wake up those nodes around event E and their opinions about the behavior of *E*. If A keeps finding that the majority of sensor nodes think that event E should not happen, A then suspects that E is malicious. After A makes a final decision, A can report this event to base stations. No matter whether it is an emergency event or a malicious event, the event can be taken care of by human operators. In practice, there may exist efficient approaches for SMM to collect information from those sensor nodes around event E. For example, Wang et al. proposed an efficient approach to construct a dominating tree to cover all the neighbors of a suspect (node B in our example). Their approach includes those nodes that have more neighbor co detectors (nodes that can provide useful information). By doing so, an efficient dominating tree can be constructed and utilized for an initiator (node A in our example) to collect information about the suspect. Krontiris et al. also proposed a voting based mechanism for collaborative intrusion detection in wireless sensor networks. In each node is equipped with a local detector module. A general algorithm consisting of Initialization Phase, Voting Phase, Publish Key Phase, Exposing the Attacker, and External Ring Reinforcement Phase is proposed to incorporate local alarms. In our future work, we plan to integrate our EKF based location module with this general algorithm, to make our system resilient to more general attacks.

Now, we present our EKF based local detection algorithm. A sensor node monitors its neighbor's behavior 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 illustrated in Algorithm 1. Here _ is a predefined threshold. In Algorithm 1, A's role is to decide whether zk+1 is abnormal or not. Node A can overhear node B's transmission zk+1 at time tk+1. After estimating ˆx+ k at time tk, A can predict node B's transmitted value ˆx −k+1 at time tk+1 based on (3). At time tk+1, A overhears B's transmitted value zk+1 and comparesˆx−k+1 with zk+1 to decide whether B is acting normally or not. If the difference between ˆx − k+1 and zk+1 (denoted as Diff in Algorithm 1) is larger than _, a predefined threshold, A then raises an alert on B. Otherwise, A thinks that B functions normally. Apparently, _ is a very important parameter here. We will provide the analysis of in Section V-B6. In practice, anomaly based IDSs suffer from a high false positive rate. We can use a post-processing scheme to reduce potential false alarms. For example, we can modify Algorithm 1 at line 4 so that *A* can raise an alert on *B* after several continuous observations of _ < Diff. The intuition here is that intrusion sessions usually demonstrate locality, i.e., many alarms within a short time window. In this way, many alerts can be used to generate one intrusion report and false alarms can be effectively reduced.

III.THERSHOLD ANALYSIS

In WSNs, various factors, such as packet loss, packet collision, time asynchrony, in aggregation protocols may contribute to uncertainties of aggregated values. Let *U* denote the variance of this uncertainty. Based on three sigma control limits in Shewchart control charts can be set to 3U. We provide the analysis of *U* in the following.

Each *Ni* represents one sensor node and each *Ni* transmits value *vi* to its parent node *N* based on a predefined aggregation protocol. Suppose that the expectation of each *vi* is $E[vi] = \mu i$ and the variance of each *vi* is $var(vi) = \sigma i\ 2$. Suppose that with a probability $0 < p < 1$ (*p* is the probability that *N* does not receive the packet from its child because of packet loss, packet collision, etc.), a packet on each link is lost. Let a random variable *X* denote the aggregated value at node *N*. We analyze the variance of *X* considering different packet loss probabilities.

IV.CONCLUSION

The project, we first proposed the integration of system monitoring modules and intrusion detection modules in the context of WSNs and then anomaly detection based aggregation

data transmission and then find the injected data for sending the malicious node. In malicious node are highlight with the red color and then collect the correct information to the original node for the data transmission in networks. find the packet loss for the Normal detection and Intruder detection

in the secure networks and then normal detection packet loss is very less and then intruder detection packet loss very high in the data transmission.

REFERENCES

[1] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Englewood Cliffs, NJ: Prentice- Hall/Simon and Schuster Company, 1993.

[2] D. Wagner, ―Resilient aggregation in sensor networks,‖ in *Proc. ACM SASN*, 2004, pp. 78–87.

[3] C. Castelluccia, E. Mykletun, and G. Tsudik, ―Efficient aggregation of encrypted data in wireless sensor networks,‖ in *Proc. MOBIQUITOUS*, Jul. 2005, pp. 109–117.

[4] H. Cam, S. Ozdemir, P. Nair, and D. Muthuavinashiappan, ―Espda: Energy efficient and secure pattern-based data aggregation for wireless sensor networks,‖ in *Proc. IEEE Sensors*, Oct. 2003, pp. 732–736.

[5] D. Chu, A. Deshpande, J. M. Hellerstein, and W. Hong, ―Approximate data collection in sensor networks using probabilistic models,‖ in *Proc. IEEE ICDE*, Apr. 2006, pp. 48–59.

[6] J.-Y. Chen, G. Pandurangan, and D. Xu, ―Robust computation of aggregates in wireless sensor networks: Distributed randomized algorithms and analysis,‖ *IEEE Trans. Parallel Distributed Syst.*, vol. 17, no. 9, pp. 987–1000, Sep. 2006.

[7] H. Chan, A. Perrig, and D. Song, ―Secure hierarchical InNetwork aggregation in sensor networks,‖ in *Proc. ACM CCS*, 2006, pp. 278– 287.

[8] K. Wu, D. Dreef, B. Sun, and Y. Xiao, ―Secure data aggregation without persistent cryptographic operations in wireless sensor networks,‖ *Elsevier Ad Hoc Networks J.*, vol. 15, no. 1, pp. 100–111, 2007.

[9] L. Hu and D. Evans, ―Secure aggregation for wireless networks,‖ in *Proc. Workshop Security Assurance Ad Hoc Netw.*, Jan. 2003, pp. 384– 391.

[10] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, ―Impact of network density on data aggregation in wireless sensor networks,‖ in *Proc. ICDCS*, 2002, pp. 457 458.

[11] S. Madden, M. J. Franklin, J. Hellerstein, and W. Hong, ―TAG: A tiny aggregation service for ad-hoc sensor networks,‖ in *Proc. OSDI*, Dec. 2002, pp. 131–146.