INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN 2278 – 0211 (Online)

An Effective Malicious Node Detection Scheme for Adaptive Data Fusion under Time-Varying Byzantine Attack

Sachin R. Gulshette
M. Tech. Student, South East Asian Institute of Technology, Karnataka, India
B. R. Prasad Babu
Professor and HOD, CSE (PG) and R & D in SEA College of Engineering and Technology, Karnataka, India

Abstract:
This paper provides details about practical idea for identification of Byzantine modernalization in communication area by involving distributed sequential network coding. The implementation complexity is minimum and the possibility of exact simulation is maximum also in the presence of huge fault nodes. Simulation results show clearly identify the faulty sensors nodes with high accuracy. In this by introducing the similarly sequential relationship in between the concept contents and the network size, we proposed a simple q-out-of-m concept which can effectively minimize the conceptual complexity, and at the instant time providing better result. We indicate that for a known package of effective sensors nodes, the identification result of the simple q-out-of-m concept enhances almost equally as the network size enhances.

We propose a simplified but efficient model to identify the effected sensors nodes before concluding the result. The result of the verified concept is evaluated under both static and dynamic attacking concepts. It is considered that with the pre-detection method, the result of the q-out- of-m concept can be enhanced significantly involved many attacked concepts. Simulation results are produce to demonstrate the effectiveness of the existing method.

Keywords: *Wireless sensor nodes, Alaram, Short Paths, False Malicious nodes*

1. Introduction

This paper explores reliable data fusion in wireless sensor network under byzantine attack in time- varying conditions by employing random linear network coding. The major limitation with it is that missed node detection and false alarm rate. We confirmed the false - alarm rate decreases as the network size increases. Since the packets are forwarded by one end hosts to other end host such network are determinable to Byzantine error nodes mentioned by compromised end hosts nodes.

Fault existing to wireless sensor nodes are regular because of sensor device itself and the hash environment where the sensor nodes are deployed, these tiny sensor nodes can easily be deployed into a designed area to form a wireless network and perform specific functions .The traditional testing and fault detection in computer systems are carried out in the form of built it Self_ Test subsystems, the major contribution of this paper in the development of a generic localized fault detection for wireless sensor networks. In this paper, we consider reliable data fusion in SENMA systems under Byzantine attacks, where a portion of the wireless nodes are compromised to report false information to the mobile access point. We propose to mitigate the Byzantine attacks using the q-out-of-m scheme, for which the final decision is depend on q sensor nodes result within that of m polled sensor nodes. However, because of the huge communication complexity provided to find the scheme parameters by required search, the optimized q-out-of-m concept could be infeasible as the network size increases. To overcome this drawback, effective sub-optimal schemes with low computational complexity are highly desired. Wireless sensor communication area are increasively embedded for set of programs such as Animal life inheritance visualization, forest fire preservation, and military surveillance for security. In these processes, the data conducted by sensor nodes from their existing environment needs to be assembled at a host computer for further analysis. Typically, an verified value is genereted at the data collector by using the corresponded verified functional working value to the encapsulated data. In huge sensor networks, calculating aggregates in network, that is mixing non completed results at intermediate nodes during message passing, significantly minimizes the amount of communication data and hence the energy utilized. An approach used by several data acquisition systems for sensor networks is to construct a spanning tree rooted at the data sink, and then perform in-network aggregation along the hierarchy. Partial results propagate level-by-level up the hierarchy, with each node awaiting messages from all its children before sending a new partial result to its parent..Scientists have designed several energy-efficient models for computing summarized using the hierarchical-based approach.

Hierarchical-based approach, however, are not robust to communication losses which result from node and transmission failures and are relatively common in sensor networks. Because each communication failure loses an entire sub-hierarchy of readings, a large fraction of sensor readings are potentially unaccounted for at the data sink, leading to a significant error in the summarized computed.

2. Related Work

In distributed systems, fault detection and identification has long being the method of active message passing research. A huge numbers of testing network connections amongst units in multiprocessor fault identification, which is too cost expensive and even not allowed in more applications at a time. They launched a common method to fault dialysis that is widely manageable and only needs a limited number of communication connections among units of that network. This model proposes a preferable vote raising among the besides of a unit to determine the status of the unit. The issue of security based network communications in the pre-presence of Byzantine attackers has been checked extensively .A survey of information theoretic research in this area is given in two important major problems are security and authentication; this work relates the later concepts. Like one-time pads, our approach relies on the generation of random values unknown to the adversary, though the one-time provides pad secrecy and not authenticity.

3. System Model

This paper use the SENMA architecture, where we assume the network is composed of n power-limited sensor nodes and a powerful mobile access point. The sensor nodes communicate directly with the mobile access point (MA), and hence no routing is required . Each sensor node detects the presence of a target object and sends its one-bit hard decision report to the mobile access point (the target is present),which makes the final decision accordingly. Under Byzantine attack, some malicious sensors send false information that would disrupt the network functionality. In our system model, we assume that the network contains k malicious sensors. The percentage of malicious sensors k/n is denoted by α and it is assumed to be known at the mobile access point. If no prior knowledge of α exists, the MA would assume a relatively high α (30%) is considered.

The main objective is to minimize the overall false alarm rate (Qf) while keeping the overall miss detection (Qm) below certain predefined value β. Hence, it is desired to get the optimum parameters m and q that can achieve the objectives. The problem can be formulated as follows:

$$\min Qf\,(m, q);$$
$$\text{s.t. } Qm(m, q) \leq \beta;$$
$$\text{s.t. } 1 \leq q \leq m \leq n;\ q,m \in \mathbb{N}.$$

In order to obtain a closed form expressions for Qf and Qm, define Pd,m−d k,n−k as the probability of polling m − d out of n − k benign sensors and d out of k malicious sensors such that:

In order to obtain a closed form expression for Qf and Qm,we define $P^{d}{}_{k,n-k}{}^{,m-d}$ as the probability of polling m-d out of n-k begin sensors and d out of k malicious sensors and d out of k malicious sensors such that:

$$P^{d,m-d}{}_{k,n-k} = \frac{\binom{k}{d}\ \binom{n-k}{m-d}}{\binom{n}{m}}$$

4. The Simplified q-out-of-m Scheme:

In the simplified approach, we set m = n, and use the following relation to obtain 'n' curve at a specific percentage of malicious nodes is the suboptimal value at a network size n1, and is the optimal value at a network size . Both and are at percent of malicious sensors.

5. Security Challenges:

Data acquisition systems for sensor networks can be classified into two broad categories on the basis of the data collection methodology employed for the application:

Query-based systems In query-based systems, the base station broadcasts a query to the network and the nodes respond with the relevant information. Messages from individual nodes are potentially aggregated en-route to the base station. Finally, the base station computes one or more aggregate values based on the messages it has received. In some applications, queries may be persistent in nature resulting in a continuous stream of data being relayed to the data sink from the nodes in the network. For such applications, the query broadcast by the base station specifies a period nodes in the network send their readings to the base station after each epoch. Event-based systems in event-based applications, such as perimeter surveillance and biological hazard detection, nodes send a message to the base station only when the target event occurs in the area of interest. If multiple reports being relayed correspond to the same event, they can be combined by an intermediate node on the route to the base station. Data acquisition systems can also be categorized based on how sensor data is aggregated. In single-aggregator approaches, aggregation is performed only at the data sink. In contrast, hierarchical aggregation approaches make use of in-network aggregation.

Hierarchical aggregation schemes can be further classified into tree-based schemes and ring- based schemes on the basis of the topology into which nodes are organized. As most existing data management and acquisition systems for sensor networks are vulnerable to security attacks launched by malicious parties. Sensor nodes are often deployed in unattended environments, so they are vulnerable to physical tampering. Since current sensor nodes lack hardware support for tamper-resistance, it is relatively easy for an adversary to compromise a node without being detected. The adversary can obtain confidential information from the compromised sensor and reprogram it with malicious code. Moreover, the attacker can replicate the compromised node and deploy the replicas at various strategic locations in the network.

A compromised node can be used to launch a variety of security attacks. These attacks include jamming at the physical or link layer as well as other resource consumption attacks at higher layers of the network software. Compromised nodes can also be used to disrupt routing protocols and topology maintenance protocols that are critical to the operation of the network. In this paper however, we focus on attacks that target the data acquisition protocol used by the application. Specifically, attacks in which the compromised nodes send malicious data in response to a query. By using a few compromised nodes to render suspect the data collected at the sink, an adversary can effectively compromise the integrity and trustworthiness of the entire sensor network. In heirarchy-based systems, compromised nodes can be used to send false event reports to the base station with goal of raising false alarms and depleting the energy resources of the nodes in the network. This paper refer to this attack as the false data injection attack to discuss an approach for filtering the false data injected by malicious nodes. Similarly, in query-based systems, compromised nodes can be used to inject false data into the network with the goal of introducing a large error in the summarized value computed at the data sink. The summarized computed by the sink is erroneous in the sense that it differs from the true value that would have been computed if there were no false data values included in the computation. Unlike event-based systems, however, in summarized systems the effectiveness of a false data injection attack depends on both the summarized value being computed, and whether sensor data is aggregated en-route to the data sink or only at the data sink, and thus the techniques used for preventing this attack differ from the techniques used in heirarchy-based systems.

6. Localized Faulty Sensor Detection

In this paper, we first give some definitions for the variables.

6.1. Definitions

We list the notations used in our algorithm and analysis below,

- n: total number of sensors;
- p: probability of failure of a sensor;
- k: number of neighbor sensors;
- S: set of all the sensors;
- N(Si): set of the neighbors of Si;
- xi: measurement of Si;

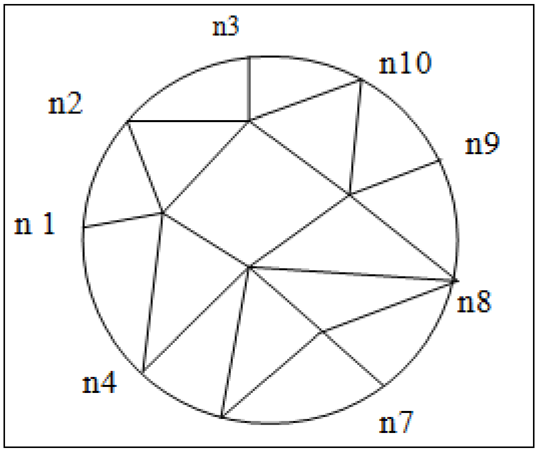

Figure 1: A partial set of sensor nodes in a wireless Sensor networks with faulty sensors

7. Simulation Results

In an effort to search for an easier and more flexible distributed data fusion solutions that can easily adopt to unpredictable environmental changes and cognitive behavior of malicious nodes we can derive a closed-form solution for the q-out-of-m fusion rule based on the central limit theorem. It is observed that the closed-form solution is a function of the network size, the percentage of malicious users, the malicious nodes behavior and the detection accuracy of the sensor nodes. We show that under a fixed percentage of malicious node, the false alarm rate for both approaches diminishes exponentially as the network size increases.

This analysis reveals an interesting and important result: even if the percentage of malicious nodes remains unchanged, large size networks are much more reliable under malicious attacks. We propose a simplified, linear q-out-of-m rule that can be easily applied to large size networks. The basic idea is to find the optimal scheme parameters at relatively small network sizes through exhaustive search, and then obtain the fusion parameters for large network size by exploiting the approximately linear relationship between the scheme parameters and the network size.

8. References

i. A. Bharathidasas and V. Anand, “Sensor networks: An overview,” Technical report, Dept. of Computer Science, University of California at Davis, 2002.

ii. C. Chong and S. Kumar, “Sensor networks: evolution, opportunities, and challenges,” Proceedings of the IEEE, vol. 91, no. 8, pp. 1247 –2056, aug. 2003.

iii. C. Karlof, N. Sastry, and D. Wagner, “Tinysec: a link layer security architecture for wireless sensor networks,” in Proceedings of the 2nd international conference on Embedded networked sensor systems, ser. SenSys ’04. New York, NY, USA: ACM, 2004, pp. 162–175. [Online]. Available: http://doi.acm.org/10.1145/1031495.1031515

iv. Ziv Bar-Yossef, S. Ravi Kumar, and D. Sivakumar. Sampling algorithms: Lower bounds and applications. Proc. of 33rd STOC, pages 266–275, 2001.

v. M. Bellare, R. Guerin, and P. Rogaway. XOR MACs: New methods for message authentication using finite pseudorandom functions. In Proc. of the 15th Annual International Cryptology Conference on in 2006

vi. L. Buttyan, P. Schaffer, and I. Vajda. RANBAR:RANSAC-based resilient aggregation in sensor networks. In Proc. of ACM Workshop on Security of Sensor and Adhoc Networks (SASN), 2006.

vii. Haowen Chan, Adrian Perrig, and Dawn Song. Secure hierarchical in-network aggregation in sensor networks. In Proceedings of ACM Conference on Computer and Communications Security (CCS), 2006.

viii. J. Considine, F. Li, G. Kollios, and J. Byers. Approximate aggregation techniques for sensor databases.2005

ix. J. Feldman, T. Malkin, C. Stein, and R. A. Servedio, “On the capacity of secure network coding,” in Proc. 42nd Annu. Allerton Conf. Communication, Control, and Computing, Monticello, IL, Sept./Oct. 2004.

x. K. Bhattad and K. R. Nayayanan, “Weakly secure network coding,” in Proc. WINMEE, RAWNET and NETCOD 2005 Workshops, Riva del Garda, Italy, Apr. 2005.