Prevention of Byzantine Attack by Large Wireless Sensor Network
C Muruganantham\textsuperscript{1} M Vanmathi\textsuperscript{2}
\textsuperscript{1}\textsuperscript{2}Department of Computer Science and Engineering

Abstract—This paper explores reliable data fusion in mobile access wireless sensor networks under Byzantine attacks. The q-out-of-m rule, which is popular in distributed detection and can achieve a good trade-off between the miss detection probability and the false alarm rate. However, a major limitation with it is that the optimal schema parameters can only be obtained through exhaustive search, making it infeasible for large networks. In this paper, first, by exploiting the linear relationship between the schema parameters and the network size, proposed simple but effective sub-optimal linear approaches. Second, for better flexibility and scalability and derive a near-optimal closed-form solution based on the central limit theorem. Third, subjecting to a miss detection constrain, prove that the false alarm rate of q-out-of-m diminishes exponentially as the network size increase, even if the percentage of malicious nodes remains fixed. Finally, propose an effective malicious node detection schema for adaptive data fusion under time-varying attacks; the proposed schema is analyzed using the entropy-based trust model, and shown to be optimal from the information theory point of view. The proposed approaches under both static and dynamic attacks.

Key words: Distributed Detection, Security In Wireless Sensor Networks, Byzantine Attacks

I. INTRODUCTION

Wireless Sensor Network has received significant attention from the research community due to their impact on both military and civilian applications. Wireless sensor networks are vulnerable to by Byzantine attacks, where the adversary has full control over some of the authentication nodes and can perform arbitrary behaviour to disrupt the system. The reliable data fusion in wireless sensor networks with mobile access points (SENMA) under static and dynamic Byzantine attacks.

In SENMA, the mobile access point traverses the network and collects the sensing information from the individual sensor nodes. In this Paper, we consider about Wireless Sensor Network in SEMA system under Byzantine attacks, where a portion of the wireless nodes are compromised to report false information to the mobile access point. We proposes to mitigate the Byzantine attacks using the q-out-of-m schema, for which the final decision is based on q sensing reports out of m polled nodes. However, due to the large computational complexity required to find the schema parameters by exhaustive search, the optimal q-out-of-m schema would be infeasible as the network size increase. To overcome this drawback, effective sub-optimal schemas with low computational complexity are highly derived.

First, we propose a simplified, linear q-out-of-m schema can be easily applied to large network. Second in an effort to search for easier and more flexible distributed data fusion solution that can easily adapt to unprintable environment changes and cognitive behaviour of malicious nodes, we derive a closed-form solution for the q-out-of-m fusion schema based on the central limit theorem. Third, we perform theoretical analysis for both linear approach & the closed-from solution. Finally, we propose a simple effective malicious node detection approach, where the malicious sensor is identified by comparing the decision of the individual sensor with that of the fusion decision.

II. SYSTEM MODEL

We 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 duality with the Mobile Access (MA), and hence no routing is required. Each sensor node detects the presence of a target object and sends its one-bit hard decision reports to the MA (1 means the target is present), which makes the final decision accordingly.

Under Byzantine attacks, 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 $\alpha$ and it is assumed to be known at the mobile access point if no prior knowledge of $\alpha$ exits, the MA would a relatively high $\alpha(30\%$ is considered high enough) and then it would obtain an-estimate for by comparing the individual sensing reports with the final decision. We also assume that each sensor has a false alarm probability $P_{f}$ and a misses detection probability $P_{m}$.

III. MODEL OF POSSIBLE ATTACK STRATEGIES

There are different attack strategies that could be adapted by the malicious sensors. Let $P_{o}$ be the probability that each malicious node intentionally reports the opposite information to its actual sensing decision. It is assumed that all malicious nodes have the same probability of attack in a particular sensing period. We classify the possible attack strategies in two categories:

- Static Attacks: In this strategy, the malicious nodes send opposite data with an arbitrary probability $P_{o}$ that is fixed, with $0<P_{o}<1$.
- Dynamic Attacks: In this strategies, the malicious node change $P_{o}$ after attacking block, which is composed of one or more sensing periods.

IV. SIMPLIFIED DATA FUSION SCHEME- THE LINEAR APPROACH

To develop effective sub-optimal schemes with low computational complexity, it is important to know how the parameters m and q change with the system variables, such as $\alpha$ and n. In this section, we consider the case where the malicious sensor attack with probability $P_{a}$. We calculate at different $P_{a}$ values under different network sizes and different percentages of malicious sensors. A simplified q-
out-of-m scheme by exploiting the linear relationship between the scheme parameter at relatively small network size.

The main idea is that we can get the optimal schema parameters at relatively small network sizes, and we them as reference points. The optimal (m, q) pairs for the different network sizes, Pa values and a ratios, can be obtained and stored in a look-up table, then used to get the suboptimal scheme parameters for large network sizes.

V. MALICIOUS NODE DETECTION AND ADAPTIVE FUSION ALGORITHM

To enhance three system performances through malicious node detection, where the hostile behaviour is identified and the malicious sensors are discarded from the final decision making. Furthermore, an adaptive fusion procedure, where the fusion parameters are tuned on the attacks behaviour and the percentage of the malicious sensors

A. Malicious Node Detection Algorithm

In the proposed approach, each node is denoted by Ti, o and present. Trustm (i) is defines in a similar way by replacing 

\[ 
\text{Trustm}(i) = \begin{cases} 
1 & \text{if decision of node } i \text{ is not equal to the final decision} \\
0 & \text{otherwise}
\end{cases} 
\]

When decision is "0" and the final decision is "1", increment Ti, o.

Hence, if \( \frac{T_{i, o}}{N} \) is greater than Pm + \( \delta \) or \( \frac{T_{i, 1}}{N} \) is greater than Pf + \( \delta \), the reports of node i, and hence m, are discarded. If decision of node i is not equal to the final decision, increment T1, 1.

VI. ADAPTIVE FUSION: CLOSED-FORM SOLUTION WITH MALICIOUS NODE DETECTION

The overall false alarm rate over observation periods when Nth=100. It can be seen that significant performance improvement is achieved for both static and dynamic attack when the adaptive fusion with malicious node detection is employed. The miss detection constraint is satisfied for all cases. The non-smoothness of the curves is mainly due to the tuning of the integer value scheme parameters to satisfy the miss threshold \( \delta_0 \), \( \delta_0 \), \( \delta_1 \) and \( \delta_2 \) have a direct impact of the performance of the malicious node detection scheme and they could be further optimized to improve the performance.

The effect of the observation interval N on the detection accuracy required longer observation interval to be detected than nodes adopting static attack. Demonstrates that schema is efficiency and highly accurate. The false alarm rate of the malicious node detection scheme (nf) is plotted versus the observation interval. As expected, that nf decrease as more observation is available at the mobile access point.

VII. ANALYSIS FROM THE ENTROPY POINT OF VIEW

In the proposed malicious node detection approach, each node’s behaviour is determined based on the uncertainty in the accuracy of its sensing report. We analyze the proposed approach using the entropy based trust model. First, for each node \( i \in \{1, 2 \ldots n\} \) we define two trust metrics \( \text{Trust}(i) \) and \( \text{Trust}(i) \) to represent the uncertainty in the node’s accuracy when the target is absent and present. \( \text{Trust}(i) \) is defines in a similar way by replacing \( P_a, f(i) = \phi \).

<table>
<thead>
<tr>
<th>Case</th>
<th>Entropy-based trust model Vs the proposed malicious node detection approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Discard suspicious report</td>
<td>Trust(( i \leq \phi )) ( \Rightarrow ) ( P_a, f(i) \geq \lambda_p, f )</td>
</tr>
<tr>
<td>1. Discard unreliable node</td>
<td>Trust(( i \leq \phi )) ( \Rightarrow ) ( P_a, f(i) \geq \lambda_p, f )</td>
</tr>
</tbody>
</table>

Table 1: Equivalence between the Entropy based trust model and the proposed malicious node detection

First, for each node \( i \in \{1, 2 \ldots n\} \) we define two trust metrics \( \text{Trust}(i) \) and \( \text{Trust}(i) \) to represent the uncertainty in the node’s accuracy when the target is absent and present. \( \text{Trust}(i) \) is defines in a similar way by replacing \( P_a, f(i) \). The entropy trust metrics are in the range \([-1, 1]\), where negative values mean that the attack probability of the corresponding node is greater than 0.5. The Trust metrics are equal to ‘1’ when the corresponding node is benign with a prefer detection accuracy. Note that \( P_f \) and \( P_m \) are generally quantities.

Our discussion above show that under the same settings for \( \delta_0,f, \delta_0,m, \delta_1, \delta_2 \), the proposed malicious node detection schemes is optimal from the information theory point of view.

VIII. CONCLUSION

An effective malicious node detection scheme for adaptive data fusion under time-varying attacks. The detection procedure is analyzed using entropy-defined trust model, and has shown to be optimal from the information point of view. It is observed that nodes launching dynamic attacks take longer time and more complex procedures to be detected as compared to these conducting static attacks. The
adaptive fusion procedure has shown to provide significant improvement in the system performance under static and dynamic attacks.

We provide a way to identify the static and dynamic attacks by calculating vulnerable points of the nodes. In our further work, we will develop a trusted network that completely deletes the malicious node from sensor networks at the moment and thereby prevent the intruder to sense the messages. We will conduct further research on adaptive detection under byzantine attacks with soft decision reports.

REFERENCE


