A Robust Iterative Filtering Technique for Data Aggregation in Wireless Sensor Networks

Nikita Pareek, Prof N. R. Wankhade
Department of Computer Engineering
KCT’s Late G.N. Sapkal College of Engineering,
Ajaneri Nashik, India
nikkipareek99@gmail.com, nileshrw_2000@yahoo.com

Abstract

Aggregation has been known to be majorly vulnerable to node compromising attacks. Since WSN are usually unattended, they are highly susceptible to such attacks. Thus, determining trustworthiness of data and reputation of sensor nodes has become crucially important for WSN. Iterative filtering algorithms seem great promise for such a purpose. Such algorithms concurrently aggregate data from multiple sources & provide trust assessment of these sources, usually in a form of similar weight factors allotted to data provided by each source. In concern with the security, we proposed improvement for iterative filtering techniques by rendering an initial approximation for such algorithms which makes them not only collusion robust, but also faster converging. It is conceived that so modified iterative filtering algorithms have a great latent for implementation in the future WSN. We extended the IF algorithms with a novel approach for collusion detection and revocation based on an initial approximation of the aggregate values as well as distribution of differences of each sensor readings.

1. Introduction

All Trust & reputation have been proposed as an efficient security mechanism for Wireless Sensor Networks (WSNs). Although sensor networks are being more and more deployed in many application domains, assessing trustworthiness of reported data from distributed sensors remains a challenging issue. Iterative Filtering (IF) algorithms are an attractive option for WSNs as they provide solution for data aggregation and data trustworthiness assessment and using a single iterative procedure. Such trustworthiness significantly differ from such estimate are assigned less trustworthiness and consequently in the aggregation process in the present iteration their readings are given a lower weight.

In recent years, several IF algorithms have been proposed for trust and reputation systems those algorithms exhibit ampler robustness compared to the simple averaging techniques; however, those algorithms have not been designed by considering more advanced collusion attack scenarios. If the attackers have knowledge about the aggregation algorithm and its parameters, attacker can launch more sophisticated attacks on WSNs by exploiting false data injection through a number of compromised nodes. In this paper we presents a new sophisticated collusion attack scenario against a number of existing IF algorithms based on false data injection.

Such initial estimation makes IF algorithms robust against such a sophisticated collusion attack, and, we believe, more robust under significantly more general circumstances; This is in contrast with the conventional non iterative statistical sample estimation technique which are not robust against false data injection by a number of compromised nodes & which can be sternly skewed in the presence of complete sensor failure. Furthermore, IF algorithms with a new approach for collusion detection and sensor node revocation is augmented. Estimate of each sensor is dependent on the distance of the readings of such a sensor from the estimate of the correct values, obtained in the preceding round of iteration by some form of aggregation of the readings of all sensors. Such aggregation is usually a weighted average.

2. Literature Survey

All S. Ozdemir and Y. Xiao[2] studies the relationship between security & data aggregation process in wireless sensor networks. Based on the current research, the open research areas and future research directions in secure data aggregation concept are provided.

Trust and reputation systems have a significant role in supporting operation of a wide range of distributed systems, from WSN & e-commerce infrastructure to social networks, by providing an assessment of trustworthiness of participants in these distributed systems. Audun Josang and Jennifer Golbeck in [3] presented Challenges for Robust Trust and Reputation Systems. This paper discusses research challenges for trust and reputation systems, and proposes a research agenda for formulating sound and authentic robustness
principles and mechanisms for trust and reputation systems.

K. Hoffman, D. Zage, and C. Nita-Rotaru [4] focus on attacks and defense mechanisms in reputation systems and provide more tight approaches to equate existing systems and to bring understanding of these systems to a broader audience, holding those who build systems that rely on reputation systems.

Provenance-based trustworthiness assessment in sensor networks [5] proposes a systematic method for assessing the trustworthiness of data items. This approach uses the data cradle as well as their values in computing trust scores, that is, quantitative measures of trustworthiness.


The second IF method conceives a correlation based ranking algorithm proposed by Zhou et al. in [7]. In this algorithm, trustworthiness of each sensor is obtained based on the correlation coefficient between the sensors readings and the current estimate of the true value of the signal.

The third algorithm considered has been proposed by Laureti et al. in [8] and is an IF algorithm based on a weighted averaging technique similar to the basic IF algorithm he only difference between these two algorithms is in the discriminant function.

[9] establish a framework to systematically tackle the challenging problem of information decoding in the presence of monolithic and surplus data. When applied to a voting system, our method simultaneously ranks the raters and the raters using only the evaluation data, consisting of an array of scores each of which represents the rating of a rater by a rater.

Ayday et al. proposed a slight different iterative algorithm in [10]. Their main differences from the rest algorithms are:

1) The ratings have a time-discount factor, so in time, their importance will fade out; and

2) The algorithm maintains a black-list of users who are especially bad raters.

The main objective of B.-C. Chen, J. Guo, B. Tseng, and J. Yang in [11] are to introduce “Bias-smoothed tensor model”, which is a Bayesian model of rather high complexity. Although the existing IF algorithms consider simple cheating behavior by opponents, none of them take into account sophisticated malicious scenarios such as collusion attacks.

Resilient Aggregation in Sensor Networks [12] paper studies security for data aggregation in sensor networks. Current aggregation schemes were planned without security in mind and there are easy attacks against them.

S. Ganeriwal, L. K. Balzano, and M. B. Srivastava in [13] proposed a general reputation framework for sensor networks in which all node employs a reputation estimation for other nodes by observing its neighbors which make a trust community for sensor nodes in the network.

Xiao et al. in [14] proposed a trust based framework which employs correlation to find faulty readings. Also, they brought in a ranking framework to associate a level of trustworthiness with each sensor node based on the number of neighboring sensor nodes are supporting the sensor.

Li et al. in [15] proposed PRESTO, model-driven predictive data management architecture for hierarchical sensor networks. PRESTO is a two tier framework for sensor data management in sensor networks. The main idea of this framework is to consider a number of proxy nodes for managing sensed data from sensor nodes.

The main contribution of authors in [16] is to propose a combination of trust mechanism, data aggregation, and fault tolerance to enhance data trustworthiness in Wireless Multimedia Sensor Networks (WMSNs) which considers both discrete and continuous data streams.

Ho et al. in [17] proposed a framework to detect compromised nodes in WSN and then enforce a software attestation for the detected nodes. They reported that the revocation of detected compromised nodes cannot be executed due to a high risk of false positive in the proposed scheme. The main idea of false aggregator detection in the scheme proposed in [18] is to employ a number of monitoring nodes which are running aggregation operations and providing a MAC value of their aggregation results as a part of MAC in the value computed by the cluster aggregator.

These studies focus on detecting false aggregation operations by an adversary, that is, on data aggregator nodes getting data from source nodes & generating wrong aggregated values. Consequently, they address neither the problem of false data being provided by the data sources nor the problem of collusion. However, when an adversary adds false data by a collusion attack scenario, it can affects the results of the honest aggregators and thus the base station will receive skewed aggregate value. In this case, the compromised nodes will attest their false data and consequently the base station assumes that all reports are from honest sensor nodes. Although the aforementioned research take into account false data injection for a number of simple attack scenarios, and work addresses issue in the case of a collusion attack by compromised nodes in a manner which employs high level knowledge about data aggregation algorithm. Although using the initial reputation results provided by our method makes IF algorithms more robust than their original version, the attacker can still alter considerably the reputation results of the IF algorithms. Thus, in this section we propose a novel attacker detection
technique in order to further diminish the impact of the compromised nodes.

3. Proposed System

3.1. Network Model

Fig. 1 depicts network model in wireless sensor network. The sensor nodes forms clusters, and each cluster has a cluster head which acts as an aggregator. At particular time interval data are collected and aggregated by the aggregator. We concentrate on algorithms which make aggregation secure when the single sensor nodes might be compromised & can send false data to the aggregator. We assume that each data aggregator has enough computational power to run an iterative filtering algorithm for data aggregation. Cluster head is selected on the basis of energy i.e. a node having highest energy is selected as a cluster head. Base station is stationary and has no energy constraint.

3.2 Robust iterative filtering for data aggregation

In order to amend the functioning of IF algorithms against the aforementioned attack scenario, our proposed approach provide a robust initial estimation of the trustworthiness of sensor nodes to be used in the first iteration of the IF algorithm. Most of the conventional statistical estimation methods for variances involve use of the sample mean. For this reason, proposing a robust variance estimation method in the case of skewed sample mean is essential part of our methodology.

Algorithmic strategy
Input : X,n,m.
Output : Reputation vector r.

1. For each sensor Si (i=1… n) do,
   check energy level
   If (Esi > Esj)
   Si is selected as CH
   Else Si acts as normal Sensor nodes.
2. Calculate the initial reputation vector using MLE
3. Calculate Weight based on distance of reading to initial reputation vector.
4. Iterative filtering with initial weights.
   a. l<=0
   b. Compute r(l+1)
   c. Compute d
   d. Compute w(l+1)
Where l is number of iteration
5. The nodes have less weight are considered as compromised.
\[ es = xs - r \]
where es is error xs sensor reading and r is the estimated reputation vector.
6. Reapply step 2 to 4 to produce more accurate readings.
7. Stop.

4. Results

Fig. 3 Impact of collusion detection module on the performance of aggregation framework.

Series 2 represent the values without collusion detection and series 3 represent the accuracy with collusion detection module. Results are calculated on the basis of how much the value of the sensor differs from the mean value. Therefore graph is plotted as standard deviation vs. RMS (Root Mean Square Error). RMS is a frequently used Measure of the Differences between values predicted by an estimator and values actually observed. Iterative filtering with and without collusion detection is shown and it clearly displays that RMS is reduced by this robust iterative filtering technique with collusion detection.

5. Conclusion

In this paper advancement in IF algorithms by providing an initial approximation of the trustworthiness of sensor nodes is provided. The IF algorithms with a novel approach for collusion detection and revocation based on an initial approximation of the aggregate values as well as differences of each sensor readings are presented. The results generated shows that the proposed approach is better than the existing system approaches.

Acknowledgement

My sincere thanks go to KCTs Late G N Sapkal college of Engineering for providing strong platform to develop skills and capabilities. I could never have completed paper without the support and assistance of many people. First and foremost, I would like to express deepest gratitude to my guide Prof. N. R. Wankhade for his excellent guidance, valuable suggestions and kind of encouragement in academics. I would also like to thank Prof. J. V. Shinde and Prof. B. R. Nandwalkar as they have given time to time information about the project. Last but not the least I owe a debt to my parents who are the silent guides in my life. I am also thankful to all my friends for their encouragement and support.

References


