ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

A Robust Data Aggregation Based Reputation Systems in Presence of Collusion Attacks for Wireless Sensor Networks

Shashi Kumara B R, Dr. Satynarayan Reddy

M. Tech Student, Dept. of Computer Science Engineering, Cambridge Institute of Technology, Bangalore, India

HOD, Dept. of Information Science Engineering, Cambridge Institute of Technology, Bangalore, India

ABSTRACT: As there is a limitation in computational power and energy resources, these two parameters becomes big challenges for WSN (Wireless Sensor Nodes). For data aggregation for sensor network use simple algorithm named as averaging due to above mentioned limitations, this simple algorithm creates many threats like malicious attacks. To overcome such threats in this paper we are using iterative filtering algorithms. These algorithms simultaneously aggregates data from multiple sources and provide trust estimation of corresponding sources in the form of weight factors assigned to data provided by each and every source. We are using Gaussian variables to estimate bias and MLE for variance estimation for reputation vector estimation. Identification of a new sophisticated collusion attack against IF based a reputation system which reveals a severe vulnerability of Iterative Filtering algorithms, deployment of an efficient and robust aggregation method inspired by the MLE, which utilizes an estimation of the noise parameters.

KEYWORDS: Gaussian variables, MLE, Robust aggregation method and Iterative filtering algorithms.

I. INTRODUCTION

Due to limitations of energy resource and computing power the data aggregation is done using simple averaging methods as a result these algorithms are more susceptible to malicious attacks and these attacks cannot be solved using cryptographic methods because when the node is compromised attacker gains access to complete data stored in that particular node. Thus for this reason we need to ascertain trustworthiness of nodes. We need more sophisticated algorithm for aggregating data. Such algorithm should produce estimates which are close to the optimal ones in presence of stochastic errors. The main target of malicious attackers is aggregation algorithms of the trust and reputation systems. Trust and reputation systems are very effective mechanism providing security for Wireless Sensor Networks (WSN's). Sensors which are in the hostile environment are susceptible to attacks where attackers inject false data into system. So, assessing trustworthiness of data and announcing it to decision makers is challenging task. UE to a need for robustness of monitoring and low cost of the nodes, wireless sensor networks (WSNs) are usually redundant. Data from multiple sensors is aggregated at an aggregator node which then forwards to the base station only the aggregate values. At present, due to limitations of the computing power and energy resource of sensor nodes, data is aggregated by extremely simple algorithms such as averaging. However, such aggregation is known to be very vulnerable to faults, and more importantly, malicious attacks. This cannot be remedied by cryptographic methods, because the attackers generally gain complete access to information stored in the compromised nodes. For that reason data aggregation at the aggregator node has to be accompanied by an assessment of trustworthiness of data from individual sensor nodes. Thus, better, more sophisticated algorithms are needed for data aggregation in the future WSN. Such an algorithm should have two features.

In the presence of stochastic errors such algorithm should produce estimates which are close to the optimal ones in information theoretic sense. Thus, for example, if the noise present in each sensor is a Gaussian independently distributed noise with zero mean, then the estimate produced by such an algorithm should have a variance close to the Cramer- Rao lower bound (CRLB), i.e., it should be close to the variance of the Maximum Likelihood Estimator (MLE). However, such estimation should be achieved without supplying to the algorithm the variances of the sensors, unavailable in practice. 2. The algorithm should also be robust in the presence of non-stochastic errors, such as faults and malicious attacks, and, besides aggregating data; such algorithm should also provide an assessment of the reliability and trustworthiness of the data received from each sensor node. Reputation and trust concepts can be used to overcome the compromised node detection and secure data aggregation problems in WSNs. The problem of collusion and the problem of false data being provided by the data sources both are not addressed in the above literature. However, when an adversary injects false data by a collusion attack scenario, it affects the results of the honest aggregators and as a result the base station will receive skewed aggregate value. In this case, false data will be attested by the compromise nodes and thus all the reports are from honest sensor nodes is assumed by base station. Even if the mentioned research considers false data injection for a number of simple attack scenarios, collusion attack scenario by the compromised nodes has not been addressed anywhere.

II. LITERATURE SURVEY

B. Awerbuch Et al. [1] proposed an iterative algorithm for trust and reputation management, a slight different iterative algorithm. Their main differences from the other algorithms are: 1) the ratings have a time-discount factor, so in time, their importance will fade out; and 2) the algorithm maintains a blacklist of users who are especially bad ratings. Although the existing IF algorithms consider simple cheating behavior by adversaries, none of them take into account sophisticated malicious scenarios such as collusion attacks. S. Ganeriwal, L Et al. [2] has proposed a general reputation framework for sensor networks in which each node develops reputation estimation for other nodes by observing its neighbors which make a trust community for sensor nodes in the network. Proposed a trust based framework which employs correlation to detect faulty readings. Moreover, they introduced 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. M. Li Et al. [3] 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 node. An interdependency relationship between network nodes and data items for assessing their trust scores based on a cyclical framework. L.-A. Tang Et.al [4] proposed a trust framework for sensor networks in cyber physical systems such as a battle-network in which the sensor nodes are employed to detect approaching enemies and send alarms to a command centre. Although fault detection problems have been addressed by applying trust and reputation systems in the above research, none of them take into account sophisticated collusion attacks scenarios in adversarial environments. Reputation and trust concepts can be used to overcome the compromised node detection and secure data aggregation problems in WSNs. S. Ozdemir Et al. [5] proposed an integration of false data detection with data aggregation and confidential transmission in wireless sensor networks. The main idea of false aggregator detection in the scheme proposed in 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. High computation and transmission cost required for MAC-based integrity checking in this scheme makes it unsuitable for deployment in WSN.

III. METHODOLOGY

Figure1 represent the overall representation of proposed system. Firstly, will consider sensors nodes with some errors and these sensors nodes are subjected to cluster formation. Once done with the selection of sensor nodes, go through all the nodes and elect cluster head from the sensor nodes having high RSS and energy to perform aggregation these cluster head acts as aggregator on the basis of CH RSS from the selected cluster clusters are formed and cluster formed nodes are subjected to robust data aggregation in this phase we are going for noise parameter estimation and Maximum likelihood estimation. Clustered nodes are modeled by using Gaussian variable in order to estimate bias by considering estimated bias value will start defining and computing variance matrices. Next, by using MLE reputation vector is estimated, the result from the estimated reputation vector is given to the Iterative filtering algorithm which reduces the number of iterations and forwards this iterations to the sink where it receives the data.

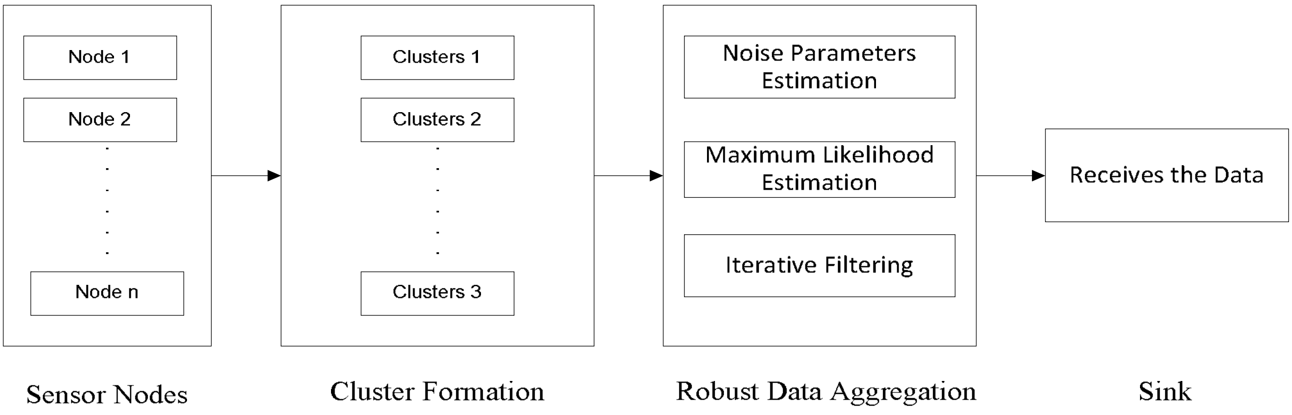


Fig. 1. Block Diagram of proposed system.

Fig. 2.

A. NETWORK MODEL WITH NODE CREATION

Firstly will consider the WSN in which sensor nodes are divided into clusters, cluster head acts as an aggregator which is present in each and every cluster of sensor node. Data are periodically collected and aggregated by the aggregator. Predicting about sensor nodes, each data aggregator has enough computational power to run an iterative filtering algorithm for data aggregation. In this method for each source in the network weighted factor is assigned. The individual id specifies the node location by allocating weight factor to each node. By assigning weight factor according to its location each and every node is identified. The allocation of weight factor is based on the computational energy need in any form of network. In this method the number of nodes connected into the network can also be identified. A common aggregation purpose is to get more information about particular groups. The network is formed and the aggregate node collects many data from multiple nodes. It is also reduce the data traffic. Once done with the selection of sensor nodes, go through all the nodes and elect cluster head from the sensor nodes having high RSS and energy to perform aggregation these cluster head acts as aggregator on the basis of CH RSS from the selected cluster clusters are formed and cluster formed nodes are subjected to robust data aggregation.

B. ROBUST DATA AGGREGATION

We need an initial estimation of the trustworthiness of sensor nodes which is to be used in the first iteration of the IF algorithm. Estimation methods for variance involve use of the sample mean. We are proposing a robust variance estimation method for skewed sample mean. Fig. 3 shows the stages of our robust aggregation framework and their connections. We provide an initial estimate of two noise parameters for sensor nodes, bias and variance in the first stage. The bias estimate is subtracted from sensors readings and in the second phase of the framework, an initial estimate of the reputation vector calculated using the MLE. In the third stage the initial reputation vector provided in the second stage is used to estimate the trustworthiness of each sensor based on the distance of sensor readings to such initial reputation vector.

It is clear that if the mean of the bias of all sensors is not zero, then there would be no way to account for it on the basis of sensor readings. On the other hand, bias of sensors, under normal circumstances, comes from imperfections in manufacture and calibration of sensors as well as from the fact that they might be deployed in places with different environmental circumstances where the sensed scalar might in fact have a slightly different value. Since by the very nature we are interested in obtaining a most reliable estimate of an average value of the variable sensed, it is reasonable to assume that the mean bias of all sensors is zero.

C. ITERATIVE FILTERING IN REPUTATION SYSTEM

We briefly describe the algorithm in the context of data aggregation in WSN and explain the vulnerability of the algorithm for a possible collusion attack. We note that our improvement is applicable to other IF algorithms as well.

Algorithm: Iterative Filtering Algorithm.

Input: X, n, m.

```
Output: The reputation vector r
Start
Step 1: Consider input X, n, m where $1 \leftarrow 0$;
Step 2: $w(0) \leftarrow 1$
Step 3: Repeat;
Step4: Compute $r\ (l + 1)$;
Step5: Compute d;
Step6: Compute$w\ (l + 1)$;
Step7: $l \leftarrow l\ +\ 1$;
Until reputation has converged
Stop
```

Where, we consider WSN with n sensors and we assume that aggregator work on one block of reading at a time each block consist of reading from m consecutive instants. Therefore the block of reading is represented by matrix X. r denotes the aggregated values it is called as a reputation vector computed with the sequence of weight w. The iterative procedure starts with giving equal credibility to all the sensors with initial value w (0). The value of the reputation vector r (l+1) in round of iteration l + 1 is obtained from the weights of the sensors obtained in the round of iteration l. The new weight vector w (l+1) to be used in round of iteration l + 1 is then computed as a function g(d) of the normalized belief divergence d is the distance between the sensor reading and reputation vector r (l). This algorithm is the iterative filtering algorithm which is vulnerable to collusion attacks improvement to this algorithm is also applicable to other IF algorithms as well. We show that the algorithm converges in little iteration.

D. ESTIMATING VARIANCE

In this section, we propose a similar method to estimate variance of the sensor noise using the estimated bias from previous section. Given the bias vector $b = [b1, b2, b3 \dots . bn]$and sensor readings$x_s^t$, we can define matrices $\hat{x}_s^t$and $\beta = \{\beta(i, j)\}$

$$\hat{x}_s^t = x_s^t - b_s \qquad (1)$$

$$\beta(i, j) = \frac{1}{m-1}\sum_{t=1}^{m}(\hat{x}_i^t - \hat{x}_j^t)^2$$

$$= \frac{1}{m-1}\sum_{t=1}^{m}\left((x_i^t - x_j^t)(b_i - b_j)\right)^2 \qquad (2)$$

By (1) we have $x_i^t - x_j^t = (r_t + e_i^t) - (r_t + e_j^t) = e_i^t - e_j^t$ thus, we get

$$\beta(i, j) = \frac{1}{m-1}\sum_{t=1}^{m}(e_i^t - b_i)^2 + \frac{1}{m-1}\sum_{t=1}^{m}(e_j^t - b_j)^2 - \frac{2}{m-1}\sum_{t=1}^{m}(e_i^t - b_i)(e_j^t - b_j) \qquad (3)$$

We assume that the sensors noise is generated by independent random variables as we have mentioned, our approximations of the bias b_i are actually approximations of the sample mean thus

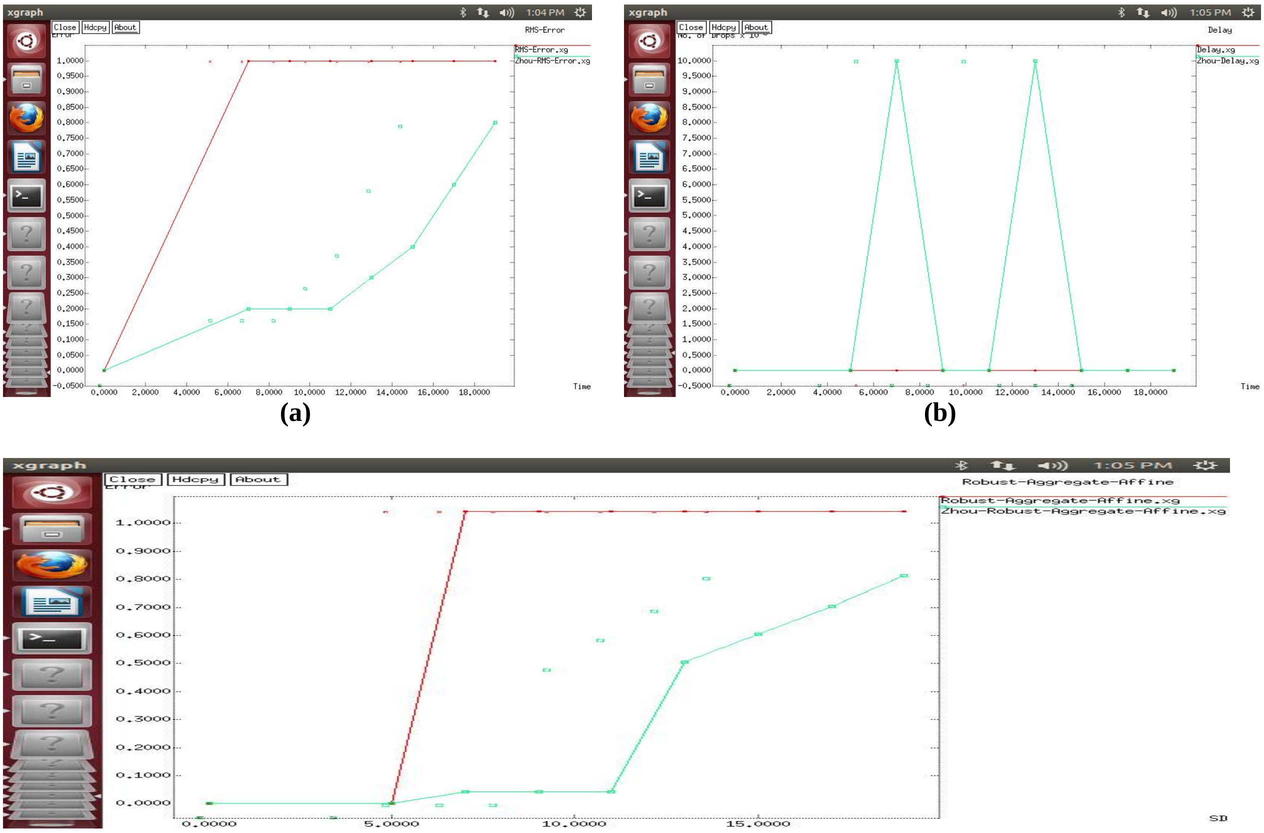
$$\frac{1}{m-1}\sum_{t=1}^{m}(e_i^t - b_i)(e_j^t - b_j) \approx \mathrm{Cov}(e_i, e_j) = 0 \qquad (4)$$

$$\beta(i, j) = \frac{1}{m-1}\sum_{t=1}^{m}(e_i^t - b_i)^2 + \frac{1}{m-1}\sum_{t=1}^{m}(e_j^t - b_j)^2 \approx \sigma_i^2 + \sigma_j^2 \qquad (5)$$

By the above equations we can prove we can estimate variance of sensor noise by computing the matrix.

IV. Experimental result

In this section, we demonstrate on a detailed numerical simulation study that examines robustness and efficiency of our data aggregation method. The main aim of our experiments is to evaluate the robustness and efficiency of our approach for estimating the true values of signal based on the sensor readings in the presence of faults and collusion attacks. For each experiment, we evaluate the accuracy based on Root Mean Squared error (RMS error) metric and efficiency based on the number of iterations needed for convergence of IF algorithms.



(a)

(b)

(c)

Figure 2: (a) RMS Matrix Comparison Graph; (b) Delay Graph; (c) Robust Aggregate Affine Graph.

V. Conclusion

In this paper we proposed architecture for estimation of sensors errors which is effective in a wide range of sensor faults and not susceptible to the described attack. In this we have introduced a sophisticated collusion attack scenario against a number of existing IF algorithms. Moreover, we have proposed an improvement for the IF algorithms by providing an initial approximation of the trustworthiness of sensor nodes which makes the algorithms robust against sophisticated collusion attacks. We designed of an efficient and robust aggregation method inspired by the MLE, which utilizes an estimate of the noise parameters obtained using contribution and enhanced IF schemes able to protect against sophisticated collusion attacks by providing an initial estimate of trustworthiness of sensors using inputs from contributions.

REFERENCES

[1] B. Awerbuch, R. Curtmola, D. Holmer, C. Nita-rotaru, and H. Rubens, “Mitigating byzantine attacks in ad hoc wireless networks,” Dept. Comput. Sci., Johns Hopkins Univ., Baltimore, MD, USA, Tech. Rep., 2004.

[2] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, “Reputation based framework for high integrity sensor networks,” ACM Trans. Sens. Netw., vol. 4, no. 3, pp. 15:1–15:37, Jun. 2008.

[3] M. Li, D. Ganesan, and P. Shenoy, “PRESTO: Feedback-driven data management in sensor networks,” in Proc. 3rd Conf. Netw. Syst. Des. Implementation, vol.3, 2006, pp. 23–23.

[4] L.-A. Tang, X. Yu, S. Kim, J. Han, C.-C. Hung, and W.-C. Peng, “Tru-Alarm: Trustworthiness analysis of sensor networks in cyber-physical systems,” in Proc. IEEE Int. Conf. Data Mining, 2010, pp. 1079–1084.

[5] S. Ozdemir and H. C¸ am, “Integration of false data detection with data aggregation and confidential transmission in wireless sensor networks,” IEEE/ACM Trans. Netw., vol. 18, no. 3, pp. 736–749, Jun. 2010.

[6] X.-Y. Xiao, W.-C. Peng, C.-C. Hung, and W.-C. Lee, “Using Sensor Ranks for in-network detection of faulty readings in wireless sensor networks,” in Proc. 6th ACM Int. Workshop Data Eng. Wireless Mobile Access, 2007, pp. 1–8.

[7] Y. Sun, H. Luo, and S. K. Das, “A trust-based framework for fault tolerant data aggregation in wireless multimedia sensor networks,” IEEE Trans. Dependable Secure Comput., vol. 9, no. 6, pp. 785–797, Nov. 2012.

[8] J.-W. Ho, M. Wright, and S. Das, “Zone Trust: Fast zone-based node compromise detection and revocation in wireless sensor networks using sequential hypothesis testing,” IEEE Trans. Dependable Secure Comput., vol. 9, no. 4, pp. 494–511, Jul./Aug. 2012.

[9] H. Chan, A. Perrig, and D. Song, “Secure hierarchical in-network aggregation in sensor networks,” in Proc. 13th ACM Conf. Comput. Commun. Security, 2006, pp. 278–287.

[10] M. C. Vuran and I. F. Akyildiz, “Spatial correlation-based collaborative medium access control in wireless sensor networks,” IEEE/ ACM Trans. Netw., vol. 14, no. 2, pp. 316–329, Apr. 2006.