IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Analysis of Secure Data Aggregation Technique for Wireless Sensor Network in the Presence of Adversary Environment based on IF algorithm

Uma Angadi[1], Dr. G.F Ali Ahammed[2]

Department of Computer Science & Engineering, VTU PG Centre, Mysuru, Karnataka, India[1]

Associate Professor, Department of Computer Science & Engineering, VTU PG Centre, Mysuru, Karnataka, India[2]

Abstract: As we have limited energy resources and computational power, data aggregation from multiple sensor nodes is done using simple methods such as averaging. WSN's are usually unattended, they are highly vulnerable to node compromising attacks. Thus making it necessary to ascertain trustworthiness of data and reputation of sensor nodes is crucial for WSN. Iterative Filtering algorithms were found out to be very helpful in this purpose. Such algorithms perform data aggregation and provide trustworthiness assessment to the nodes in the form of weight factors.These algorithms simultaneously aggregate data from multiple sources and provide a trust estimation of these sources, usually in a form of corresponding weight factors assigned to data provided by each source. In this paper we analyzed some secure data aggregation mechanisms and introduced a new complicated collision attack with its impact on wireless sensor networks.

Index Terms: Collusion Attacks,data aggregation,Iterative Filtering Algorithm,wireless sensor network.

I. INTRODUCTION

Wireless sensor networks are being increasingly deployed in many application areas, however computational power and energy resources are two big challenges for Wireless sensor networks.Their limitations cause sensor network to use a simple algorithm called averaging for data aggregation. Data aggregation using simple averaging scheme is more exposed to faults and malicious attacks. An attacker can capture and compromise sensor nodes and launch a variety of attacks by controlling compromised nodes.This cannot be prevented by cryptographic methods, because the attackers generally gain complete access to information stored in the compromised nodes. To protect against this threat, it is important to establish trust levels for sensor nodes and adjust node trustworthiness scores.Trust and reputation systems have an important role in supporting the operation of a wide range of distributed systems, from wireless sensor networks to social networks, by providing an estimation of trustworthiness of participants in such distributed systems. An estimation of trustworthiness at any given instant represents an aggregate of the behavior of the participants up to that instant and has to be robust in the presence of various types of faults and malicious behavior. There are a number of ways for attackers to manipulate the trust and reputation scores of participants in a distributed system, and such manipulation can usually harm the performance of the system. Iterative Filtering (IF) algorithms are an efficient and reliable option for wireless sensor networks because they solve both problems of data aggregation and data trustworthiness estimation using a single iterative procedure. As soon as the computational power of very low power processors significantly improves, future aggregator nodes will be capable of performing more difficult data aggregation algorithms, thus making wireless sensor networks less vulnerable.

II. RELATED WORK

Our work is related to three areas that are studied from the other reference such as IF algorithms, trust and reputation system for WSNs and secure data aggregation with compromised node detection in WSNs.There are many past literatures introduce the IF algorthims for solving data aggregation problem.In one of the prior work six different IF algorithms are proposed.They are all iterative and are similar to one another.The only difference is their choice of norm and aggregation function.A bias-smoothed tensor model based on a Bayesian model is introduced in another paper.The sensor complexity in this model is high due to its mathematical framework.The Existing filtering technique considered only the cheating behaviour of adversaries, none of them take into account of collusion attack.Our work is also related to the trust and reputation systems in WSNs. PRESTO is a model driven predictive data management architecture proposed for hierarchical sensor networks.It is a two tier framework for data management in the networks. This includes a number of proxy nodes for managing sensor readings from corresponding sensor nodes.Trust and Reputation concepts can be used to overcome the compromised node and secure data aggregation problems in sensor nodes.A combination of trust mechanisms,data aggregation and fault tolerance is also proposed to enhance data trustworthiness.It considers both discrete and continuous

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

data stream. A trust framework is proposed for sensor networks in cyber physical system such as bottle-network.In this work sensor nodes are employed to detect approaching attackersand send alarms to a command center.The main goal of false aggregator detection is to employ a number of monitoring nodes which are performs only aggregation operations and provides a MAC value of their results as a part of the value computed in the aggregator.Some of the prior works focus on detecting false aggregation on the cluster head.That is, data aggregator node obtaining data from source nodes and producing wrong aggregated values.The problem of false data being provided by the data soureces and collusion attackis not addressed in these works.

III. METHODOLOGY

A. NETWORK MODEL

The conceptual model proposed by Wagner in [4] is considered for sensor network topology. Fig. 1 shows assumption for network model in WSN. The sensor nodes are divided into seperate clusters, and each cluster has a cluster head which acts as an aggregator. Data are periodically collected and aggregated by the aggregator. Authors in [5] assume that the aggregator itself is not compromised and concentrate on algo-rithms which make aggregation secure when the individual sensor nodes might be compromised and might be sending false data to the aggregator. It also assume that each data ag-gregator has enough computational power to run an suitable algorithm for data aggregation.

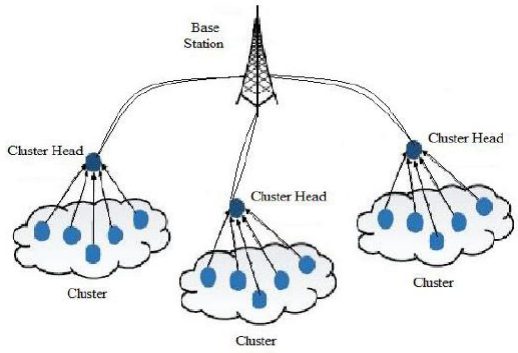

Fig .1.Network model of wirelss sensor network.

We provide a thorough empirical evaluation of effectiveness and efficiency of our proposed aggregation method. The results show that our method provides both higher accuracy and better collusion resistance than the existing methods.

B. ADVERSARY MODEL

The past researchers [1][6] developes the attack models by considering the fact that they cannot rely on cryptographic methods forpreventing the attacks, since the adversary may extract cryptographic keys from the compromised nodes. The authors in, considers Byzantine attack model, where the ad-versary can compromise a set of sensor nodes and insert any false data through the compromised nodes [7]. Following are some assumptions made in this model

a. Sensors are deployed in a hostile unattended environment with some physically compromised nodes.
b. When a sensor node is compromised, all the information which is inside the node becomes accessible by the adversary. System cannot depend on cryptographic methods for preventing the attacks because the adversary may extract cryptographic keys from the compromised nodes [8].
c. Through the compromised sensor nodes the adversary can send false data to the aggregator with a purpose of changing the aggregate values.
d. All compromised nodes can be under control of a single adversary or a colluding group of adversaries, enabling them to launch a sophisticated attack.
e. The adversary has enough knowledge about the aggregation algorithm and its parameters. The base station and aggregator nodes cannot be compromised by adversary node.

C. COLLUSION ATTACK SCENARIO

In this scenario ten sensors are assuming that report the values of temperature, which are aggregated using a suitable aggregation algorithm Most of the algorithms employ simple assumptions about the initial values of weights for sensors [9]. In the suitable adversary model, an attacker is able to mislead the aggregation system through careful selection of reported data values. The collusion attack scenarios are as follows

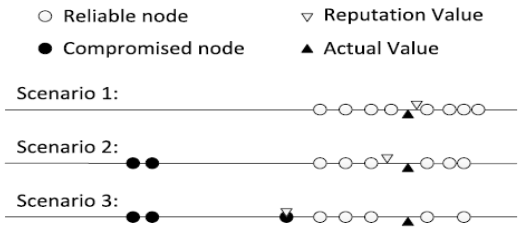

Fig.2.Collusion attack scenario.

Most of the IF algorithms occupy simple assumptions about the initial values of weights for sensors. In case of our opponent model, an attacker is able to misinform the aggregation system from side to side cautious range of report data standards. Assume that ten sensors report the values of temperature, which are aggregated using the IF algorithm planned in with the reciprocal discriminated function.

In scenario 1, all sensors are reliable and the result of the IF algorithm is close to the actual value.

In scenario 2, an adversary compromises two sensor nodes, and alters the readings of these values such that the simple average of all sensor readings is skewed towards a lower value. As these two sensor nodes report a lower value, IF algorithm penalizes them and assigns to them

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

lower weights, because their values are far from the values of the sensors. The algorithm assigns very low weights to these two sensor nodes and consequently their contributions decrease.

In scenario 3, an adversary employs three compromised nodes in order to launch a collusion attack. It listens to the reports of sensors in the network and instructs the two compromised sensor nodes to report values far from the true value of the measured quantity. It then computes the skewed values of the simple average of all sensor readings and commands the third compromised sensor to report such skewed average as its readings.

IV. PROPOSED SYSTEM ARCHITECURE

In the wireless sensor network consists of sensor nodes these sensor nodes are scattered then deployed environment in the network and then to form the cluster ,each cluster has a cluster head and then data send to the aggregator node before sending base station to verify the data, if any, error in the data, then to estimate the value using parameters such as bias and variance and also estimate MLE using an iterative filtering algorithm.The proposed system architecture view can be shown in Fig 3.

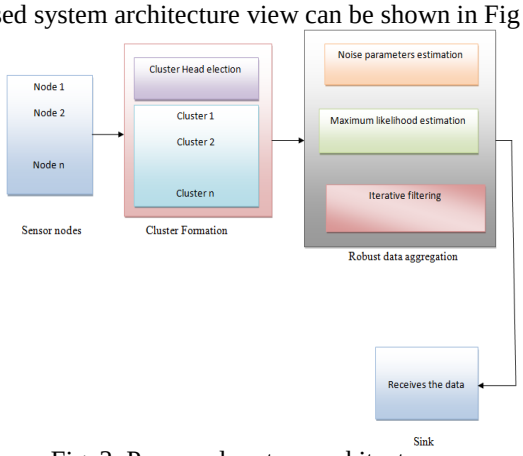

Fig .3. Proposed system architecture

V. ROBUST DATA AGGREGATION FRAMEWORK

Robust Data Aggregation model operates on batches of consecutive readings of sensors, proceeding in several stages. In the first stage provide an initial estimate of two noise parameters for sensor nodes, bias and variance details of the computations for estimating bias and variance of sensors. A novel approach for estimating the bias and variance of noise for sensors based on their readings. The variance and the bias of a sensor noise can be interpreted as the distance measures of the sensor readings to the true value of the signal. In fact, the distance measures obtained as our estimates of the bias and variances of sensors also make sense for non-stochastic errors. Based on such an estimation of the bias and variance of each sensor, the bias estimate is subtracted from sensor readings and in the next phase of the proposed framework, we provide an initial estimate of the reputation vector calculated using the MLE as shown in Fig 4..

A. Bias Estimation

All sensors may have some errors in their readings. Such error is denoted as e_s^t Of sensor is and it is modelled by the Gaussian distribution random variable with bias b_s And variance σ_s. Let r_s Denotes the true value of the sensor at time t. Sensor readings x_s^t can be written as

$$x_s^t = r_s + e_s^t \qquad (1)$$

Since there is no true value, the error value of sensors is not to be found. But the difference values of such sensors are calculated with the equation given below. Let $\delta = \delta(i,j)$be an estimator for mutual difference of sensor bias.

$$\delta(i,j) = \frac{1}{m}\sum_{i=0}^{m}(e_i^t - e_j^t) = \frac{1}{m}\sum_{t=1}^{m} e_i^t = \frac{1}{m}\sum_{t=1}^{m} e_j^t \qquad (2)$$

Let $a_i = \frac{1}{m}\sum_{t=1}^{m} e_i^t$ be the sample mean of the random variable and m be the number of readings for each sensor.Then the expected value is calculated by minimizing the obtained value with respect to the mean value and the equation is given below

$$\delta(i,j) = a_i - a_j \approx b_i - b_j \qquad (3)$$

B. Variance Estimation

With the known values of bias estimated from the equation 3 the variance of sensor errors are calculated.Each sensor bias value is substracted from the sensor readings.By using the error difference value from the equation 2 we can get the variance value as a squared difference of each sensor error and the bias value.This varies upto the last sensor reading and is defined as

$$\beta(i,j) = \frac{1}{m-1}\sum_{t=0}^{m}(e_i^t - b_i)^2 + \frac{1}{m-1}\sum_{t=0}^{m}(e_i^t - b_i)^2 \qquad (4)$$

Where $\sigma_i^2 = v_i$ is the variance of sensor from the matrix $\beta = \{\beta(i,j)\}$.

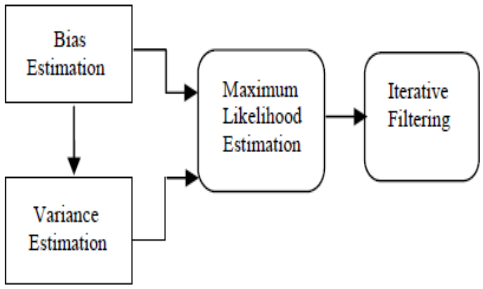

Fig.4.Data Aggregation Framework

C. Maximum Likelihood Estimation

The unbiasing sensor readings are extracted and take place with help of the bias estimated result which is calculated from the above section.After that the variance estimated result from equation 4 is considered.and the extracted unbiasing sensor is used to make the maximum likelihood estimation with variance value By differentiating the likelihood function the true values are obtained and are measured in the form of weighted average.

DOI 10.17148/IJARCCE.2016.5489

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

It is defined as

$$r = \sum_{s=1}^{n} w_s x_s \qquad (5)$$

Thus it estimates the reputation vector without any iteration. Hence the computational complexity of the estimation is less than the existing IF algorithms.

D. Enhanced Iterative Filtering

For the proposed collusion attack the results from the above is considered as an initial reputation for this filtering. It estimates the trustworthiness of each sensor based on the distance of sensors readings.

From this process the estimation is made with an initial level itself. Using this initial reputation the efficiency of the IF algorithm is improved and reduces the required number of iterations.

IF algorithm is robust against the simple outlier injection by the compromised nodes. An adversary employs three compromised nodes in order to launch a collusion attack. It listens to the reports of sensors in the network and instructs the two compromised sensor nodes to report values far from the true value of the measured quantity.

It then computes the skewed values of the simple average of all sensor readings and commands the third compromised sensor to report such skewed average as its readings.

VI. RESULTS

The objective 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.

Apply dKVD- Reciprocal, dKVDAffine, Zhou, Laureti and robust aggregation approach tosynthetically generated information. Although simply apply our robust framework to all existing IF approaches, in this paper investigate the improvement which addition of our initial trustworthiness assessment method produces on the robustness of dKVD-Reciprocal and dKVD-Affine methods.

The main shortcoming of the IF algorithms in the proposed attack scenario is that they quickly converge to the sample mean in the presence of the attack scenario. In order to investigate the shortcoming, we conducted an experiment by increasing the sensor variances as well as the number of colluders.

Fig5 represents the node placement in the network.the nodes are grouped into the cluster format.each cluster has cluster head is called aggregator node.

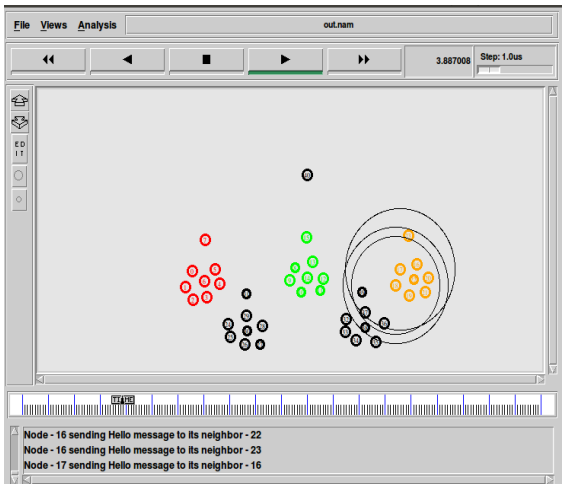

Fig 5.Placing the nodes in the network.

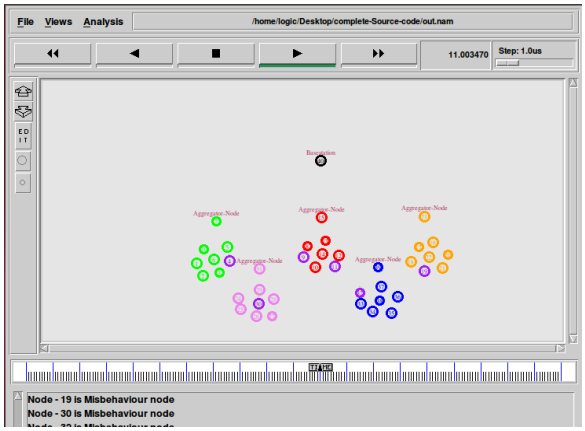

Fig 6.Data aggregation process in network.

Fig.6 represents the node are deployed in the network.to form a cluster ,in the network cluster head are 41 node is the base station and the cluster heads are 7,15,23,31 and 39.to form the aggregation process

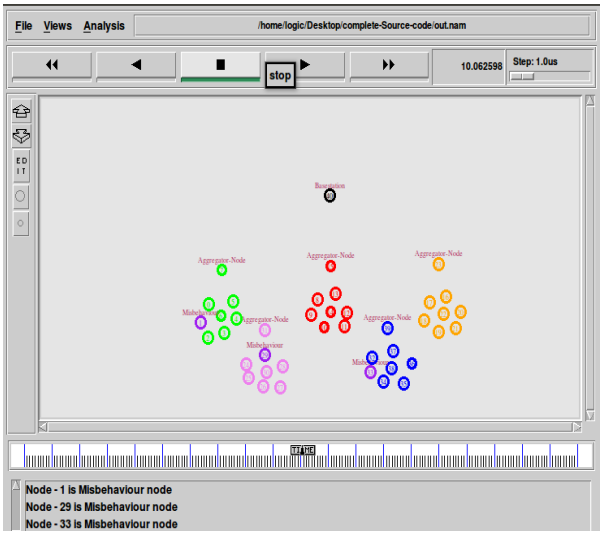

Fig.7.Node 1,29 and 33 are the malicious node.

Fig 7 represents the data aggregation In network to transmitted the data from multiple source to base station

DOI 10.17148/IJARCCE.2016.5489

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

during the aggregation process any malicious node occurs then to indentify and detect the malicious node in that process.they nodes are 1,29,33 these are the malicious nodes.

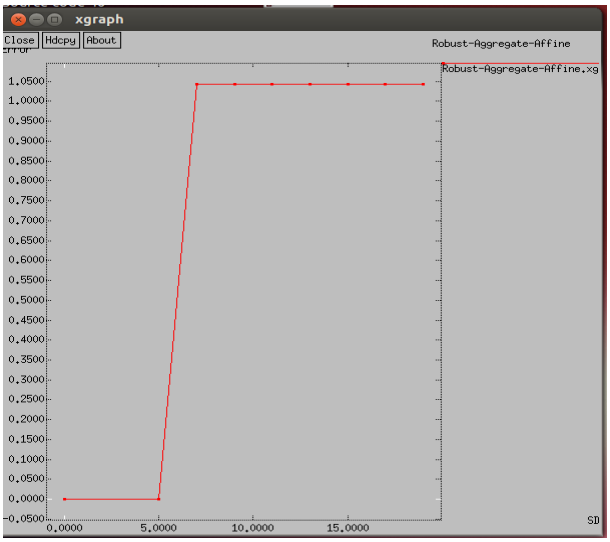

Fig 8. Robust –Aggregate-Affine.

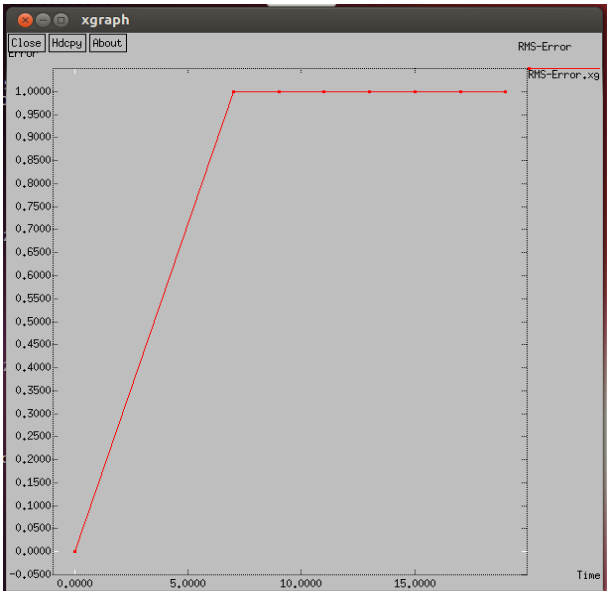

Fig 9 .RMS-Error.

These two graphs getting from above this experiment.above Fig 8 shows the robust aggregate affine graph and above Fig 9 shows the RMS-Error graph. The results obtain from this experiment show that the original version of the IF algorithm quickly converges to the skewed values provided by one of the attackers, while starting with an initial reputation provided by our approach, the algorithms require around 29 iterations, and, instead of converging to the skewed value provided by one of the attackers, it provide a reasonable accuracy. The results of this experiment show that the proposed initial reputation for the IF algorithm improve the efficiency of the algorithm in terms of the number of iterations until the process has converged. This can be explained by the fact that the new initial reputation is close to the true value of signal and the IF algorithm needs fewer iterations to reach its stationary point.

VII. CONCLUSION AND FUTURE WORK

Data aggregation mechanisms along with data averaging techniques are analysed. Network model proposed by Wagner is described for sensor network network. Adversary models with their assumptions are reviewed. New sophisticated collu-sion attack scenarios along with its impact on wireless sensor networks is explained. As soon as computational power of very low power processors significantly improves, future ag-gregator nodes will be capable of performing more difficult data aggregation algorithms, thus making wireless sensor networks less vulnerable. In future an enhanced strategy against collusion attack is introduced which makes is not only collusion robust, but also more accurate and faster converging.

REFERENCES

[1]. Mohsen Rezvani, Aleksandar Ignjatovic, Elisa Bertino, and Sanjay Jha, “Secure Data Aggregation Technique for Wireless Sensor Networks in the Presence of Collusion Attacks” , IEEE Transactions on Dependa-ble and Secure Computing (TDSC) ,2014

[2]. Y. Sun, H. Luo, and S. K. Das, “A trust-based framework for fault-tolerant data aggregation in wireless multimedia sensor networks” , IEEE Transaction on Dependable & Secure Computing ,Nov. 2012.

[3]. H.-S. Lim, G. Ghinita, E. Bertino, and M. Kantarcioglu, “A game theoretic approach for high-assurance of data trustworthiness in sensor networks ” , IEEE International Conference on Data Engineering (ICDE), April 2012.

[4]. D. Wagner, “Resilient aggregation in sensor networks,” in Proc. 2nd ACM Workshop Security Ad Hoc Sens. Netw., 2004, pp. 78–87.

[5]. Y. Yang, X. Wang, S. Zhu, and G. Cao, “SDAP: a secure hopby-hop data aggregation protocol for sensor networks,” in MobiHoc, 2006, pp. 356–367.

[6]. E. Ayday, H. Lee, and F. Fekri, “An iterative algorithm for trust and reputation management,” in Proceedings of the 2009 IEEE international conference on Symposium on Information, 2009.

[7]. B. Awerbuch, R. Curtmola, D. Holmer, C. Nita-rotaru, and H. Ru-bens, “Mitigating byzantine attacks in ad hoc wireless networks,” Department of Computer Science, Johns Hopkins University, Tech. sRep., 2007.

[8]. J. Bahi, C. Guyeux, and A. Makhoul, “Efficient and robust secure aggregation of encrypted data in sensor networks,” in Fourth Inter-national Conference on Sensor Technologies and Applications, July 2010.

[9]. 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 ” ,IEEE International Conference on Data Mining , 2010.