COMPUTERS & SECURITY 113 (2022) 102540

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

TC 11 Briefing Papers

Identifying malicious nodes in wireless sensor networks based on correlation detection

Yingxu Lai[a,b,*], *Liyao Tong*[a], *Jing Liu*[a], *Yipeng Wang*[a], *Tong Tang*[a], *Zijian Zhao*[a], *Hua Qin*[a]

[a] *Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China*
[b] *Engineering Research Center of Intelligent Perception and Autonomous Control, Ministry of Education, Beijing 100124, China*

ARTICLE INFO

Article history:
Received 20 April 2021
Revised 19 September 2021
Accepted 7 November 2021
Available online 13 November 2021

Keywords:
Autoregressive integrated moving-average model
Correlation coefficient
Correlation theory
Dempster–Shafe evidential reasoning
False data injection attacks
Wireless sensor network

ABSTRACT

The wireless sensor network (WSN) is a multi-hop wireless network that comprises multiple sensor nodes arranged in a self-organized manner. It is usually deployed in unattended areas where sensor nodes can easily be infiltrated by attackers who can affect the detection results by injecting false data. This paper proposes a malicious-node identification method based on correlation theory that prevents fault data injection attacks. First, anomalies among similar types of sensor data are detected based on time correlation. Second, malicious nodes are identified based on spatial correlation. Third, the identified malicious nodes are verified based on event correlation. The experimental results and their comparison with those of existing methods show that the proposed scheme has better recall with lower false-positive and false-negative rates than those of the traditional fuzzy reputation model and weighted-trust-based methods.

1. Introduction

Wireless sensor networks (WSNs) integrate many advanced technologies (e.g., microelectronics, embedded tools, modern networks, and wireless communications). WSNs have been widely applied to environmental monitoring (Martin et al., 2015; Tamandani and Bokhari, 2016), scientific observation (Spachos and Hatzinakos, 2015; Wang et al., 2015), and traffic monitoring (Gao et al., 2019) tasks. A WSN comprises many nodes that are low-cost and small-volume. These nodes can be generally divided into three types: sensing, cluster head, and gateway. A sensing node collects environmental information in the area of coverage and processes it through the cluster head. The gateway collects the processed information and provides responses.

In WSNs, environments can be monitored via the cooperation between nodes, which are usually placed in an uncontrolled external environment. Hence, energy, computing resources, storage capacity, communications, and other material properties are limited. Therefore, they are vulnerable to environmental factors and malicious attacks. A WSN divides abnormal nodes into two types according to anomalous

WSN, Wireless sensor network.

* Corresponding author.
E-mail address: laiyingxu@bjut.edu.cn (Y. Lai).

https://doi.org/10.1016/j.cose.2021.102540

cause. Nodes that work abnormally owing to failure or mishap are called "faulty nodes". The other type is called "malicious nodes," which refer to those that are compromised by attackers. These attacks include challenge collapsar, distributed denial of service, false data injection (FDI), and witch attacks, among others.

This paper focuses on malicious FDI attacks, wherein attackers manipulate node detection and system decisions by injecting incorrect data. Therefore, it is very important to detect these attacks in WSNs. For this reason, this paper proposes a method of identifying FDI attacks in WSNs using features of correlation theory, because the data collected by sensing nodes are related to temporal, spatial, and event-based correlations. Temporal correlations are those in which sensing data (e.g., environmental information) generally has long correlations; hence, past sensing data can affect future and present interpretations. Spatial correlations are those in which there is similarity among the data collected by adjacent sensing nodes in the monitoring area. Event-based correlations are those in which the occurrence of an event (e.g., fire) causes the nodes at the event location to change in terms of correlation capability.

This study uses temporal, spatial, and event-based correlations to detect and prevent malicious nodes in WSNs. First, based on the principle of temporal correlation, the sensing node establishes a local fault data detection mechanism on the collected environmental information using divided difference filtering (DDF) (Nørgaard et al., 2000) to identify abnormal data. Second, the cluster head leverages data uploaded by the sensing nodes to detect malicious nodes based on the principle of spatial correlation. Third, the gateway uses the information uploaded by cluster heads to verify malicious nodes based on event correlation to improve detection accuracy.

The contributions of this paper can be summarized as follows:

- To improve the accuracy of detecting anomalies, the DDF-2 algorithm is used to construct the iterative state estimation model of the prediction model, and the predicted value of the model is modified and evaluated.
- To solve the problem of inconsistent event judgment when detecting anomalies on neighbor nodes, this paper improves the method of allocating attribute weights to the AdaBoost algorithm. Additionally, a method to calculate correlations is designed, and the max–min approach is used to calculate a correlated fused value of nodes. This helps weaken the influence of a fault node on malicious-node detection and improves detection accuracy.
- An event correlation method is proposed to accurately determine malicious nodes, which considers the advantages of temporal and spatial correlations and verifies the identified malicious nodes in the WSN by locating events and its developing trend.

The remainder of this paper is arranged as follows: the current status of research in the area is summarized in Section II, and Section III introduces the FDI attack model. Section IV introduces our proposed method. Section V describes the experimental results, and an interpretive discussion is presented in Section VI. Section VII presents the conclusions of this paper.

2. Related works

The mobile edge computing sensor network is a promising service provision platform for all types of users (Liu et al., 2020). WSN node intrusion detection is an example service. Several researchers have studied WSNs in light of common attacks on the routing layer, such as denial of service, black-hole, forged routing, and routing table overflow attacks. However, malicious-node attacks are destructive to the data collected, which affects detection results. Therefore, a great deal of attention has been paid to WSN attack detection. Recent works on node intrusion detection are divided into three categories: credit-, reputation-, and acknowledgement-based systems.

The main idea of credit-based systems is reward node cooperation. Nodes receive credits for providing services to other nodes, such as helping them with packet forwarding. Nobahary and Babaie (2018) proposed a credit-based method of clustering network nodes to vote on the identification of potential malicious nodes.

In reputation-based systems, each node randomly listens to messages from its neighbors to generate corresponding degrees of trust for their forwarding behaviors. For example, Karthigadevi et al. (2019) proposed a sinkhole detection mechanism using a network density estimation technique that leveraged neighbor details to estimate the network density and identify the presence of malicious nodes in the region. Acknowledgement-based systems send acknowledgement packets in the opposite direction to confirm their correctness when receiving or sending packets.

2.1. *Detection schemes against WSN threats*

Owing to the limited resources of WSNs, secure routing is a difficult challenge. Previous research organized WSN threats into black-hole, Sybil, FDI, random poisoning, and sinkhole attacks. We discuss those in this subsection.

Kalkha et al. (2018) proposed a hidden Markov model (HMM) that identifies malicious WSN nodes to prevent black-hole attacks. This method is based on end-to-end delays and packet-delivery rates and uses HMM's empirical measurement method. Alternatively, Hammamouche et al. (2018) solved black-hole attacks using node reputation and multiple verification techniques. The value of node reputation is determined by situation and observation conditions. This method detects and eliminates simple and cooperative black-hole attackers while strengthening cooperation among nodes. In most studies, the energy consumption of trust acquisition and diffusion was very large, negatively affecting network lifetime.

With the Sybil attack, the attacker analyzes related costs and benefits before launching the attack. Kumar and Bhuyan (2020) proposed a game-theory method to resist Sybil attacks. According to its characteristics, the Kumar mechanism sets a global trust threshold to maintain node persistence in the network while identifying its trust level, making the attack expensive. Additionally, the author defined the utility functions of attacker and defender.

FDI attacks inject false data into nodes to control their decision outputs, which then affects the judgment of the entire

system. Owing to the mobility of internet-of-things (IoT) devices, it is more difficult to detect these kinds of malicious attacks in WSNs. Hence, Yaseen et al. (2018) proposed an IoT-device-tracking model based on fog computing to detect malicious attacks. Compared with the traditional WSN model, this model reduces the computing cost of aggregation nodes to a greater extent.

Hua et al. (2020) studied three different distributed delayed least-mean-square algorithms based on the Kullback–Liebler divergence (DLMSKLs) to weaken the impact of FDI attacks and analyzed their performance. The simulation results showed that the most effective DLMSKL algorithm was the Durbin–Watson (DW) version. Compared with other anti-FDI attack algorithms, the DLMSKL-DW algorithm weakens the impact of FDI attack more effectively.

Random poisoning attacks can change node behaviors to prevent detection by the system, hence attack uncertainty must be accounted for. Meng et al. (2016) proposed a high-level conspiracy attack that invades cooperative intrusion detection networks using random poisoning attacks. Meng et al., showed that the trust mechanism must account for the assumption that malicious nodes randomly send malicious feedback. They can also send false information without significantly reducing their trust values. This poses additional, more subtle challenges. It further indicates that a combination of detection techniques is necessary to defend against more sophisticated attacks.

Sinkhole attacks target IoT networks, and they are easy to launch and difficult to defend against. Liu et al. (2018) proposed a pit-attack defense scheme based on a detection route to resist these attacks. The defense utilizes a routing mechanism that combines reverse, equal-hop and minimal-hop routing of remote sinks, effectively avoiding the attack mode to find a safe route to the real sinks. The scheme utilizes the characteristics of network energy consumption, and the detection path mainly occurs in an area where remaining energy exists. Therefore, the proposed scheme has little influence on network lifetime.

2.2. *Detection schemes based on different WSN dimensions*

The methods discussed in this paper are based on node behavior assessment through which researchers have examined the detection of malicious nodes by using cooperative intrusion detection systems. However, the combination of their variety is insufficient. To determine the malicious nodes and make full use of feature information, it is instead necessary to identify malicious nodes from within temporal, spatial, and event dimensions.

Kumar et al. (2021) proposed a detection scheme based on spatial dimensions and event procedures. The scheme improved a deep convolutional neural network designed to find and isolate malicious nodes during the detection phase. The scheme clustered trusted nodes in the energy-efficient phase to achieve a balance between security and energy-savings.

Bhuiyan and Wu (2016) proposed a temporal and spatial detection scheme that identified malicious-node attacks in time and space. In the time dimension, attack-related behaviors were determined by detecting mutations in correlation. In the spatial dimension, differences in the behaviors of nodes were determined by detecting correlations among multiple nodes.

Li et al. (2020) proposed an event-based scheme that included a distributed filtering algorithm for linear systems under unknown input and FDI attacks and an event-based distributed estimator to ensure satisfactory estimation performance, even under attacks. The estimator uses the state estimation information from target and adjacent nodes to obtain the filter gain by minimizing the upper bound of the estimation error covariance.

Although several studies have focused on identifying malicious nodes, most schemes based on trust do not consider the reliable evaluation value of sensor nodes. To overcome this limitation, efforts to achieve accurate trust values have become a new driving force. Yao et al. (2014) proposed a multi-level fuzzy trust model that considers the multi-dimensionality of trust to improve its accuracy and reduce the number of rules. Jaint et al. (2018) proposed a weighted-trust method that frequently assigns trust values to nodes. With adaptable weight as the evaluation value, this method reduces erroneous evaluations. However, there is still plenty of room for improvement in both accuracy and computation cost.

Other schemes combine a variety of network intrusion detection mechanisms but do not consider how different dimensions of sensor-node data can achieve satisfactory performance. Our scheme combines the temporal, spatial, and event-based dimensions and their correlations to improve malicious-node identification. In our experiment, we compare our method to a traditional credit-based method (Nobahary and Babaie, 2018), two weighted-trust methods (Yao et al., 2014; Jaint et al., 2018), and the most recent FDI attack (Tufail et al., 2021) to prove our method's effectiveness.

3. FDI attack definition

When a FDI attack occurs, the attacker attempts to take control of the system or influence its judgment by injecting spurious information into the network (Xie et al., 2010). In WSNs, the nodes are vulnerable to FDIs because they are usually positioned in uncontrolled environments. An attacker tampers with a node's sensing data to force the greater system to make an incorrect conclusion about the current environment. For example, with a WSN used for fire monitoring, temperature, humidity, and smoke information can be manipulated to create a false positive. Additionally, such an attack can cause a false negative in which a real fire is not detected. The false scenario generated by this attack is constructed with Eq. (1):

$$D_{\mathrm{False}} = D_{real} + \alpha, \tag{1}$$

where, α represents the data inserted by an attacker, D_{real} are the real data, and D_{False} are the falsified environmental information caused by the injection attack. The FDI attack assumes that the attacker only attacks the sensing node, and the cluster head and gateway are not subject to FDI. In reality, the cluster head and gateway nodes are strongly protected. The goal of attackers is usually to cause the entire WSN to misjudge the current environment. Hence, the attack behavior includes tampering with data of abnormal or normal states.

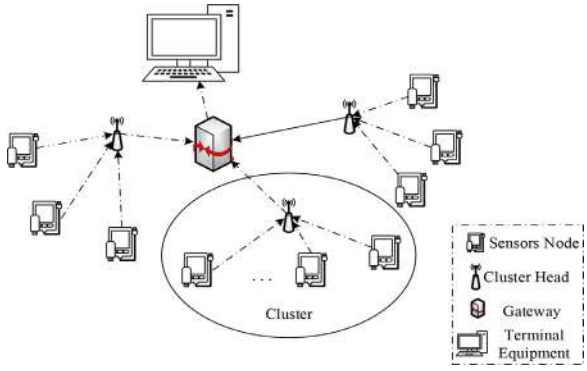


Fig. 1 – Distributed detection mode.

4. Correlation theory applied to malicious-node detection

This paper leverages the correlation between nodes to detect malicious ones. We now describe methods of detection and principles of malicious nodes that drive the FDI attacks.

There are two ways to detect malicious nodes: centralized and distributed means. Centralized detection relies on all sensing nodes that transmit data to the gateway, and they are summarized and analyzed. The centralized detection method consumes a large amount of energy as all data are uploaded. Distributed detection requires the assignment of detection tasks to the sensing nodes, cluster heads, and gateways to understand the hierarchical processing of data. Therefore, the distributed detection method consumes less energy and is suitable for detecting FDI attacks. The detection model is shown in Fig. 1. The sensing node collects environmental information, and the cluster head performs the relevant calculations on the collected data. The cluster head then uploads the calculation results to the gateway, which responds accordingly.

Based on this method of detection, we can detect malicious nodes using the principle of correlation detection. First, anomalies in data from sensors are detected based on temporal correlations. In a WSN, an anomalous event's impact on a single node will be reflected in the time series of the node's perception data. The working state of the node can then be judged by observing the historical data of the node alongside the changes in the perceived data at the next time slot. Second, spatial correlations can be used to detect malicious nodes. Because environmental events that cause changes to sensing nodes in adjacent areas that are highly similar, the spatial correlation between nodes can be used to determine the working states of nodes. Finally, event correlation can be used to verify malicious nodes. Because the occurrence of events (e.g., a fire) leads to changes in node correlation, event correlations can help verify malicious nodes in only the first two steps.

Regarding the FDI model proposed in this paper, the correlation theory combines the above correlations to detect malicious nodes, as shown in Fig. 2. As stated in Section III, the attacker's purpose is to influence the entire system to make an incorrect judgment about the current environment. The adversary injects incorrect environmental data, potentially affecting all sensing nodes in the system. From the correlation between nodes, those with low correlation can be identified so that abnormal nodes can be detected. A few problems remain in the detection of each correlation and are discussed below.

4.1. *Detection of abnormal data based on temporal correlations*

This study uses temporal correlations to detect abnormalities among nodes. Owing to the long correlation of the information collected by sensing nodes, the normal range of data at the next time can be estimated by modeling the historical data. If the data collected by the sensing node exceed the normal range, they are determined to be in an abnormal node. This method accuracy mainly depends on the prediction model. However, the sensing node is vulnerable to interference from external factors that can lead to fluctuations in the collected data, such as temperature and humidity, which may lead to inaccurate predictions. The DDF algorithm can reduce the uncertainty of predictions by calculating the prior estimation of the node state and the Kalman gain of the collected data. Here, the autoregressive integrated moving average (ARIMA) modeling method is adopted to establish a mathematical model that can accurately reflect the dynamic dependencies contained in the sequence according to the limited size of the collected data. This helps us predict the data at next time slot. Specifically, the DDF-2 algorithm is used to construct an iterative state estimation model to modify and evaluate the predictive state value of the ARIMA model. The sensing node constructs the next normal range of the collected data by the predicted data state and preset threshold. By constantly updating the above steps on the sensing node, an iterative detection mechanism for erroneous data based on a threshold is formed for the correction and evaluation of the predicted state of the model. By using the collected data and the posterior estimation to correct the direction of the evolution of the system, the estimated error can be reduced.

4.2. *Detection of malicious nodes based on spatial correlation*

Malicious nodes are detected based on the spatial correlation between them. If there is an FDI attack in the WSN, the correlation fusion value of the attacked node in the cluster head will deviate from those of nearby cluster heads, helping identify malicious nodes. We detect malicious nodes by analyzing their abnormal conditions and calculating their associated fusion value. First, the cluster head uses Dempster–Shafe (D–S) evidential reasoning to fuse the abnormal conditions of each attribute datum to obtain the abnormal conditions of the nodes. D–S evidence theory is a generalization of Bayesian reasoning, which does not need prior knowledge and can deal with uncertain problems. It can express the uncertainty of research problems through upper and lower probabilities. Thus, it is widely used to deal with uncertainty problems. Second, the cluster head calculates the similarity between nodes based on each abnormal condition to obtain the associated fusion value. The malicious nodes are detected by observing the changes in correlation fusion values between

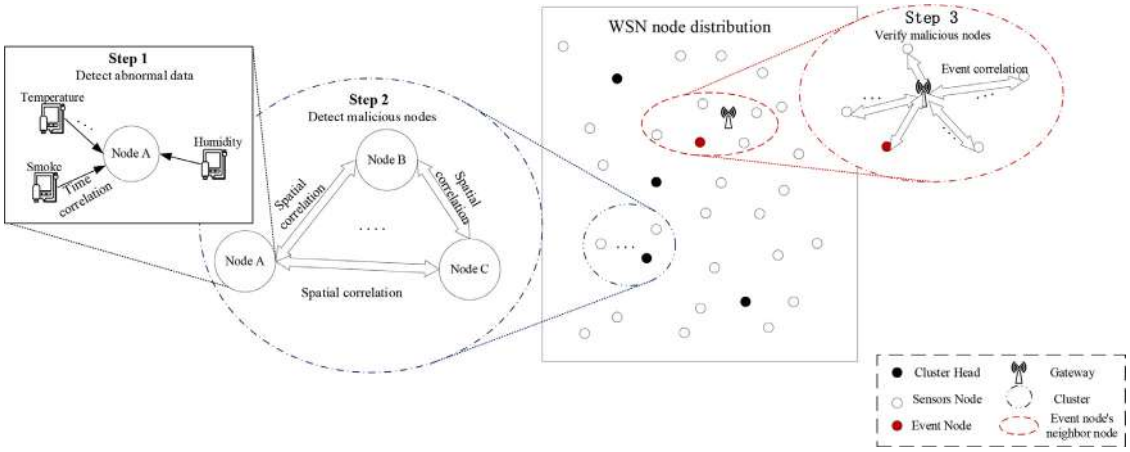


Fig. 2 – Detection of malicious nodes based on the correlation.

neighboring nodes, but each detection algorithm has limitations, which are described next.

The detection of exceptions in single-attribute data in the prior stage is often insufficient for detecting node exceptions. According to D–S evidence theory, the cluster head can fuse the abnormal situation of multi-source data to determine whether a node is abnormal. However, it encounters the problem of conflicting evidence (i.e., the Zadeh paradox) when the cluster head uses D–S evidence theory for information fusion. This paradox occurs when 100% of the trust is assigned to a proposition having a small possibility, which produces a result contrary to intuition. With actual data processing, evidence conflicts such as this are often encountered; hence, we should avoid the error to avoid reaching false conclusions, which seriously affects the accuracy of the judgment results. This study uses the AdaBoost learning algorithm to deal with such conflicts arising from D–S evidence. Because AdaBoost has a strong ability to fit and classify data, it classifies the data and processes conflicting evidence by combining weak learners to form a strong learner. Therefore, conflicting evidence is regarded as difficult to categorize and classify, owing to noise and faulty data.

WSN nodes may be affected by the detection of malicious nodes caused by node-fault transmissions because of interference from the environment. In this paper, a method to calculate correlation coefficients and weights is designed to solve these problems. The method is characterized by its severe punishment of abnormal-node occurrences, where the punishment is reduced in case of a fault. Thus, the influence of faulty nodes on detection can be reduced. In real monitoring scenarios, the method used to calculate the correlation coefficient changes with the monitoring data in each period. To optimize the weight of the correlation coefficient between nodes in the comprehensive judgment, nearness degree theory is used. Different nodes within the same time around the same target are correlated to find relevant central values. If there is a malicious attack behavior among nodes, the target object will be recognized, owing to its obvious deviation from the central value. Hence, the attack behavior will be given a small weight, which effectively weakens the impact of malicious attacks.

4.3. *Verification of malicious nodes based on event correlation*

Because temporal and spatial correlation theories do not consider trends of variation in data from the perspective of events, misjudgments or missing judgments will occur. Based on event correlation, taking a fire event as an example, the gateway can determine the location of the fire based on information of faulty nodes and can observe the development trend of the event by the temperature field near the fire to verify malicious nodes detected in the first two steps.

5. Implementation of proposed method

Malicious nodes can be detected adequately based on correlation theory because various problems occur in the data processing algorithm referenced by each correlation module. This section discusses the inaccurate prediction model in terms of temporal correlation, the inability of D–S to deal with a conflicting evidence, the calculation of the association fusion values of nodes in terms of spatial correlation, and the identification of malicious nodes in verifying event correlation. The corresponding solutions are also presented.

5.1. *DDF algorithm modifies prediction model based on temporal correlation*

The DDF-2 algorithm is introduced in this paper to construct an iterative state estimation model of the prediction model to modify and evaluate the predicted state values of the ARIMA model. The estimation error is reduced because the collected data and the posterior estimation are used to correct the direction of the evolution of the system.

The pseudocode for the detection of abnormal data based on temporal correlation is as follows:

5.2. *Spatial correlation—Improved detection algorithm based on near and long-distance correlations*

This paper improves D–S evidential reasoning in terms of near spatial correlation and the algorithm to determine the association fusion value of nodes in remote spatial correlations.

5.2.1. *Improved D–S evidence theory based on AdaBoost*

When cluster heads use D–S for information fusion, they encounter the Zadeh paradox, which seriously negatively affects the accuracy of judgment results.

For each node, sample data X (sequence of abnormal situations in the previous stage) contains T sets of behavioral data; each set has n attributes, and the weights are ρ_n, $(n = 1, 2, \cdots, N)$. Suppose that in the initial situation, the weight of each set of data is $\omega_i^1 = 1/T$, and the change in the weight of the sample after $p$iterations is ω_i^p, $(i = 1, 2, \cdots, T;\ p = 1, 2, \cdots, P)$.

Therefore, the sensor can generate Tsets of behavioral data, and each set has n attributes.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N-1} & x_{1N} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x_{i1} & \cdots & x_{in} & \cdots & x_{iN} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x_{T1} & x_{T2} & \cdots & x_{TN-1} & x_{TN} \end{bmatrix}. \tag{2}$$

In Eq. (2), each column of the matrix represents an attribute, and each row represents the abnormality of each attribute at the given moment. Therefore, the values in the matrix represent whether the attribute data are abnormal. Abnormality is represented by "0," and the normal is "1."

Using the weighted average method shown in Eq. (3), the attribute weight of each sample, ρ_n, is multiplied by the weight of the set of data, ω_i^p, and the sample data, x_{in}, under the given number of iterations. It is then divided by a normalization factor, which is a sample score vector: the gravity center of x_{in}. It is the feature representation of each sample attribute in the next set of data in the $p-th$ iteration.

$$\overline{x_n^p} = \sum_{i=1}^{T} \rho_n \omega_i^p x_{in} / \sum_{i=1}^{T} \omega_i^p, \tag{3}$$

where $n = 1, 2, \cdots, N$, $p = 1, 2, \cdots, P$.

According to the Euclidean distance, after two adjacent $p$and $p+1$ iterations, the offset in the gravity center of the sample, d_{p+1}, is

$$d_{p+1} = \sqrt{\sum_{i=1}^{N} \left(\overline{x_i^{p+1}} - \overline{x_i^p} \right)^2}, \tag{4}$$

where $p = 1, 2, \cdots, P-1$.

Eq. (4) expresses the difference between the same set of data under two iterations.

Let d be the threshold of the Euclidean distance.

1) If $d_{p+1} > d$, there is conflicting evidence, and the modified AdaBoost algorithm is used to judge it.

Let l be the correction factor. Then,

$$l = \begin{cases} \frac{d_{p+1}-d}{d_{p+1}+d}, & d_{p+1} > d \\ 1, & d_{p+1} \le d \end{cases}. \tag{5}$$

As shown in Eq. (5), l increases as distance d_{p+1} increases because d is constant, indicating that the weight assigned to the "difficult" data in the $p-th$ iteration is extremely large. Thus, the weight of the $p-th$ observational data should be adjusted according to l.

The equation for the AdaBoost-adjusted weight is Eqn 6

$$\omega_i^{p+1} = \omega_i^p \alpha_p^{1-lg_p(x_i)y_i}. \tag{6}$$

The index in this equation, α_p, is modified, and the weight is determined by l. If the value is large, the weight, ω_i^{p+1}, will reduce in the next iteration, meaning that the conflict in this iteration is large; the conflicting data should be weakened, and the impact of the conflicting evidence should be reduced.

The strong classifier is Eqn 7

$$g(x) = \arg\max_{y \in Y} \sum_{p=1}^{P} \log \frac{1}{\alpha_p} g(x_i), \tag{7}$$

where $g(x_i)$ is a weak classifier for x_i.

The AdaBoost algorithm ensures that the accuracy of most classifiers is greater than half to guarantee that the final classification error of the combined classifier set tends to zero. Therefore, when e=0or $e > 1/2$, the iterations should be stopped.

The abovementioned method uses the correction coefficient, l, to reduce the weight of conflicting evidence so that some extremely conflicting evidence is obscured and does not affect the overall judgment of the AdaBoost model. This reduces overfitting and leads to a more accurate classification.

2) When $d_{p+1} < d$, there is no conflicting evidence, and the basic probability assignment of each attribute is calculated directly by the D–S method.

The improved algorithm for node association fusion value is based on nearness degree theory. Whether a node is malicious depends on the evaluation of other nodes in the given cluster. Thus, the evaluation of node j is based on the fusion of its associated value by other nodes in the cluster. (shown in Eqn. 8

$$\Re_{ch,j}(\Delta t) = \sum_{k=1}^{n-2} \omega_k G_{k,j}, \tag{8}$$

where ω_k is the weight of the correlation between nodes. $G_{i,j}$ is the correlation coefficient between nodes, and $\Re_{ch,j}(\Delta t)$ is the associated fusion value of the given node.

The methods to calculate the correlation coefficient, $G_{i,j}$, and weight, ω_k, are designed as follows. The method to calculate the correlation coefficient is shown in Eq. (9).

$$G_{i,j}(\Delta t) = \left[\frac{10 \times r_{ij}(\Delta t)}{r_{ij}(\Delta t) + u_{ij}(\Delta t)} \left(\frac{1}{r_{ij}(\Delta t)\sqrt{u_{ij}(\Delta t)}} \right) \right], \tag{9}$$

where $r_{ij}(\Delta t)$ is the number of normal data in a period of time, and $u_{ij}(\Delta t)$ is the number of abnormal data in a period of time. As the number of exceptions increases, $1/\sqrt[r_{ij}(\Delta t)]{u_{ij}(\Delta t)}$ can gradually tend to zero and a small number of error data transmissions of nodes will not greatly impact detection. Therefore, the above method reduces the impact of fault nodes on detection.

To optimize the weight of the correlation coefficient between nodes in a comprehensive decision, this paper introduces the proximity degree theory. The maximum and minimum criteria are introduced to measure the similarity between each evaluation node, which is in the same monitoring period. Therefore, the weight of node j is evaluated using all of the information of the correlation proximity degree between node j and other nodes, as shown in Eq. (10):

$$\omega_j = \frac{\sum_{l=1}^{n-2} \sigma_{j,l}}{\sum_{j=1}^{n-2}\sum_{l=1}^{n-2} \sigma_{j,l}}, \tag{10}$$

where $\sigma_{k,l} = \min\{G_{k,j}, G_{l,j}\}/\max\{G_{k,j}, G_{l,j}\}$, and$\omega_j$ satisfies $\sum_{j=1}^{n-2} \omega_j = 1(0 \le \omega_j \le 1)$.

The average value, $\overline{\tau}$, and the standard deviation, o, of all $\Re_{ch,j}(\Delta t)$ is obtained to calculate the minimum threshold, $\overline{\tau} - o$. If the associated value of node fusion is less than the minimum threshold, the node is judged to be malicious.

The method to calculate the similarity between nodes in the cluster is designed to reduce the impact of node faults on detection, and the proximity degree is introduced to optimize the weight of the correlation coefficient between nodes in the decision and improve the detection of malicious nodes.

5.2.2. *Verifying malicious nodes based on event correlation*

Taking a fire event as an example, we locate the fire location and malicious nodes by the detection results of the above scheme and verify the findings by combining them with information about the surrounding temperature field.

Regarding the abnormal event location, if the Euclidean distance between nodes is less than their sensing radii, the two nodes are neighboring nodes in the WSN. If node S_m is a neighbor of S_k, and S_m is a normal node, then S_k is a faulty node and can be used as the dividing point between the fire area and the area with no fire. The area formed by multiple dividing points is where the fire occurs.

1) Calculate the center coordinates of the coverage area according to the dividing point as shown in Eqn. 11:

$$(x_c, y_c) = \left(\frac{\sum_{l=1}^{M} x_l}{M}, \frac{\sum_{l=1}^{M} y_l}{M}\right). \tag{11}$$

2) Calculate the distance between the center node and each dividing point and use the maximum distance as the radius as shown in Eqn. 12.

$$r = \max(d_{(x_l, y_l)}) l \in (1, 2, \cdots, M). \tag{12}$$

Therefore, the fire zone is a circular area with (x_c, y_c) as the center and r as the radius.

Kriging interpolation is used to establish the temperature field, which can be used to determine the range of the fire.

Table 1 – Experiment parameters.

Parameter	Value
Experiment Location	Huahai Grassland in Qinghai Province
Number of nodes	600
Temperature Threshold	5 °C
Humidity Threshold	10%
Smoke Threshold	30 ppm
Transmission Technology	ZigBee
Transmission distance	200 m

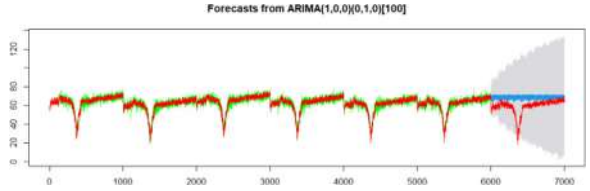


Fig. 3 – Prediction results of DDF-2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

For $S_k \in S_i, i = \{1, 2, \cdots, n\}$, if node S_k detects the fire, its temperature and that of its neighbor nodes should be high. If the above correlation is not satisfied, then node S_k is judged to be a malicious node.

6. Experimental results

First, 600 sensing nodes are distributed in the experimental environment to collect environmental information, including temperature, humidity, and smoke concentrations. The experimental environment is set to have malicious attacks at the same time of fire. Supposing that among the 600 scattered sensing nodes, nodes R4–R9 are nodes in the event area, and there are some malicious nodes in this group. Each sensor node collects 1000 groups of environmental information. A group of data is collected at each time point, t, and every 10 time points is a time period, T. Assuming that the time of the fire event is in time 300–400, malicious WSN nodes are detected through the above detection steps. The specific experimental parameters are shown in TABLE 1.

According to the actual situation of the grassland, an ARIMA prediction model was established based on the humidity of node R7, as shown in Fig. 3.

In Fig. 3, the green line represents the fitted value, the blue line represents the predicted value, the red line represents the actual value, and the gray shows the predicted confidence interval.

Analyzing the first 1000 sets of data of node R7 and in periods of time T, the ARIMA prediction model is built. In Fig. 3, a large difference is observed between the predicted value and the actual value from times 6000–7000, which shows that the humidity of node R7 was abnormal at this time period. This situation is consistent with the experimental hypothesis, which demonstrated that the ARIMA model improved with

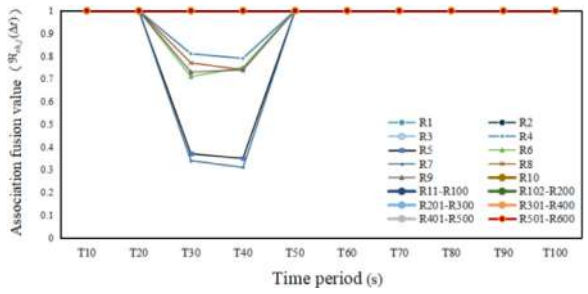


Fig. 4 – Malicious-node detection results.

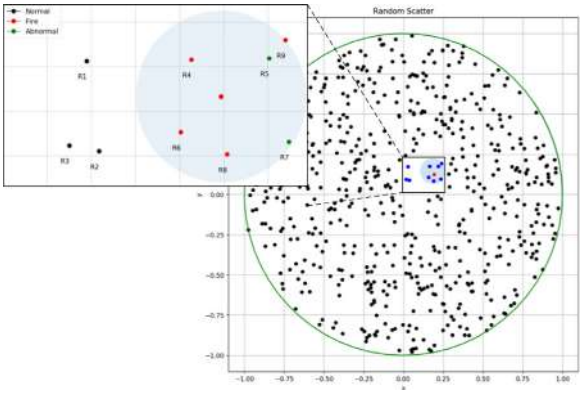


Fig. 5 – Event location.

the DDF-2 algorithm could adequately detect the abnormal nodes.

Under the premise of the above experimental scenario and prediction model, the detection of malicious nodes is realized, and the detection results are shown in Fig. 4. It can be seen from the figure that the correlation value of nodes R4–R9 differs from other nodes. Thus, it can be preliminarily judged that there may be some accidents in the region of nodes R4–R9. Moreover, through the above calculation, it can be seen that in the accident area, the association fusion values of nodes R5 and R7 are quite different from those of others. Therefore, nodes R5 and R7 can be initially determined as malicious nodes.

These results can be represented as a plane, as shown in Fig. 5. The maximum distance between the center of the circle and each dividing point was taken as the radius of the region, and the range of coordinates of the fire area was determined according to Eq. (10). The black dots in Fig. 5 are normal nodes, the blue area is the fire area, and the red dots in the fire area are nodes around the fire event. The red pentagon located in the fire area shows the location of the fire source. As shown in Fig. 5, R4–R9 are in the fire area, and R1–R3 were in the non-fire area. The results of detection are consistent with the experimental hypothesis. Therefore, the proposed algorithm is accurate in terms of calculating the scope of the fire.

We then used Kriging interpolation to establish the temperature field. Through this information, the correctness of the fire-related information and the scope of the fire site were determined. Kriging interpolation was used to transform the discrete points into a continuous 3-dimensional plane. The

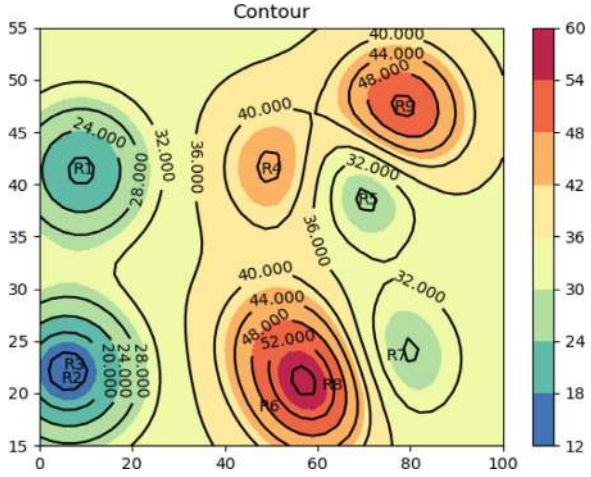


Fig. 6 – Temperature field.

results are shown in Fig. 6, from which we can see the temperature range of the region. The light-blue color represents the region where nodes R5 and R7 were located, which had a low temperature. The red color shows the region where nodes R4, R6, R8, and R9 were located, which had a high temperature. Nodes R4–R9 were in the fire area, but the temperatures of R5 and R7 did not increase significantly. Therefore, R5 and R7 were verified as malicious nodes.

7. Performance analysis

To better illustrate the advantages of the proposed method, ours with and without improvement were compared to each other, to the fuzzy trust algorithm proposed by Nobahary and Babaie (2018), and to the credit-based method proposed by Yao et al. (2014). Assuming that there are FDI attacks in the environment, the ratio of malicious nodes to all nodes was taken as the experimental variable. Recall, false-positive rate (FPR), and false-negative rate (FNR) were used to evaluate the effectiveness and reliability of the system. The evaluation equations are as followsEqn 13, 14 and 15:

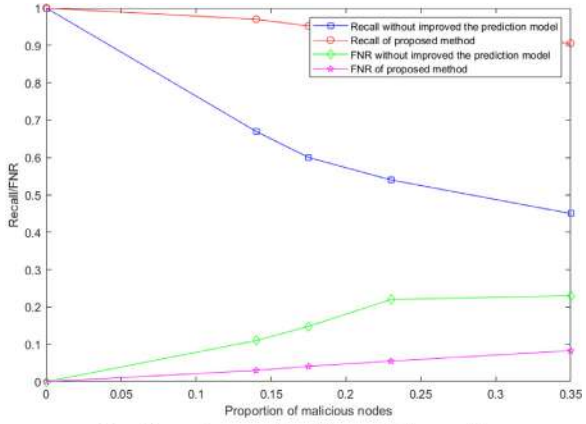
$$Recall = \frac{TP}{TP + FN}, \tag{13}$$

$$FPR = \frac{FP}{TN + FP}, \tag{14}$$

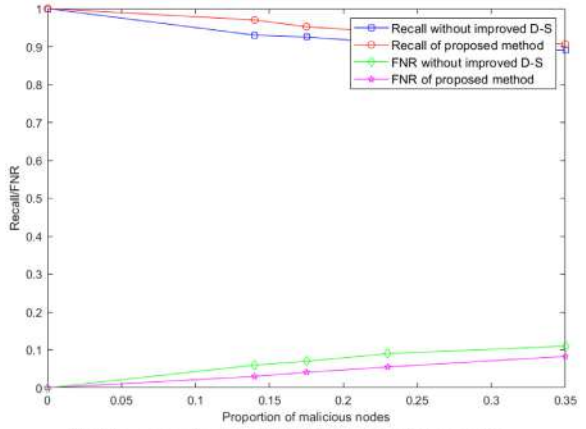
$$FNR = \frac{FN}{FN + TN}, \tag{15}$$

where TP is the number of true positive samples that are classified as positive, FP is the number of negative samples that are classified as positive, TN is the number of negative samples classified as negative, and FN is the number of positive samples classified as negative.

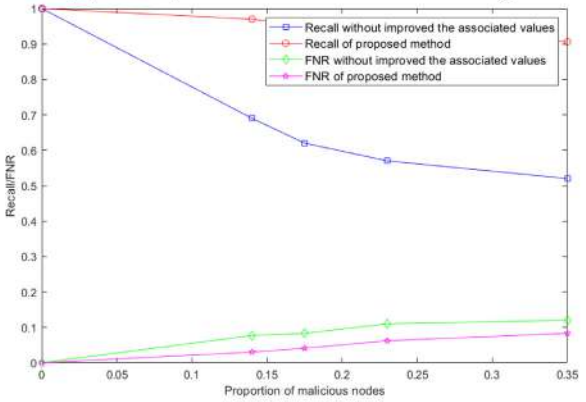
Our method with and without improvements in Recall and FNR are shown in Fig. 7. Fig. 7(a) represents the results of the comparison, and Fig. 7(b) represents those of D–S evidential reasoning with and without improvements. Fig. 7(c) shows



(a) Comparison results of the prediction model

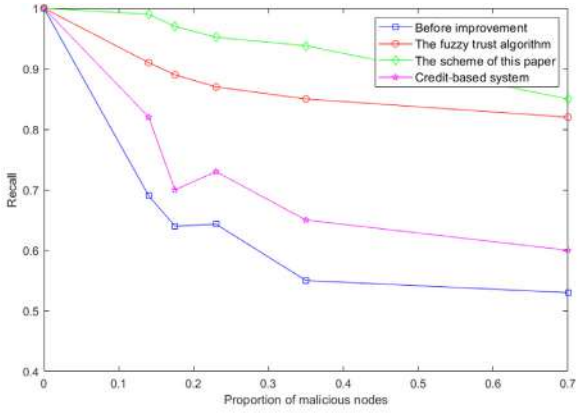


(b) The comparison results of D-S evidential reasoning

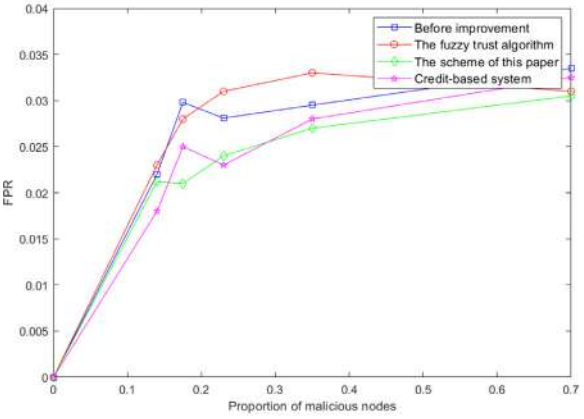


(c) The comparison results of associated fusion values

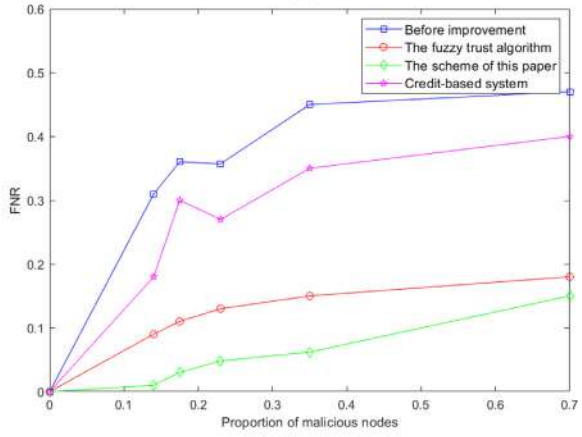
Fig. 7 – Comparison results with and without improvement.



(a)



(b)



(c)

Fig.8 – Our method's comparison results vs. other methods.

the results with and without the improvement to the method used to calculate the associated fusion values. They all show that improvements in the prediction model and values of correlation fusion have a significant impact on the experimental results, whereas the improvement in D–S evidential reasoning has a smaller impact.

According to the results in Fig. 8, the recall of six methods showed a downward trend with the increase in the ratio of malicious nodes, where our method lacking improvements recorded the steepest decline. Compared with the methods proposed by Tufail et al. (2021), Yao et al. (2014), and Jaint et al. (2018), and Nobahary and Babaie (2018), our method based on correlation theory was superior because it has the advantages of correlation detection for malicious nodes. This can be seen from Fig. 8(b) and (c), in which the FPR and FNR detected by malicious nodes are in positive proportion to the proportion of malicious nodes. Under the same ratio of malicious nodes, the proposed detection method based on correlation exhibited lower FPR and FNR than those of the other five methods.

The correlation among nodes was used to detect malicious nodes. The FPR of the proposed scheme was lower than that of the other five methods, and all were lower than 3%. The FNR was less than 30%, which is lower than that of other methods. This shows that the proposed method yielded efficient performance in terms of detecting malicious nodes based on correlation.

8. Conclusion

This paper proposed a method to detect FDI attacks based on correlation analysis. In terms of time correlation, owing to the instability of data collected by nodes, the DDF-2 algorithm was used to modify the prediction model, effectively reducing the estimation error over previous methods. Because the fault transmission of nodes and the attack-related behavior of malicious nodes affect detection rates in WSNs, we designed two malicious-node detection methods. First, the cluster head uses the detection results of the previous stage and improves the AdaBoost algorithm to optimize the weight allocation of each attribute by comprehensively considering the abnormal situation of each attribute data. The algorithm also reduces the impact of node-fault transmission on detection. We also designed a method to calculate the correlation coefficient and used the max–min approach to calculate the correlation fusion values of nodes. This is helpful in weakening the influence of the fault node on the detection of malicious nodes and improve the accuracy of detection.

An event correlation-based detection method was proposed to improve detection performance. A fire event was used as an example. By locating the event and utilizing information on the temperature field around the event node, we identified and verified malicious nodes in the WSN. The comparative analysis of experiments found that both the fuzzy and weighted-trust algorithms must calculate the coefficient according to a manually configured weight by subjective consciousness, resulting in a lack of stability and reliability of trust evaluation results. The credit-based methods did not consider the complexity of the relevant indicators, which limited their applicability. Artificial-intelligence-based method considered nodes in isolation, which made it impossible to cope with changes of node data caused by the current environmental changes. This reflects the advantage of detecting malicious nodes based on correlation theory. The experimental results showed that the proposed method in this paper is superior to other methods in recall, FPR, and FNR. Therefore, the scheme proposed in this paper shows good performance.

Although the proposed method is better than the compared method according to the above three indicators, the proposed method also has some limitations. For example, with the increase of the proportion of malicious nodes, the recall of the proposed method has a downward trend. When the proportion of malicious nodes is 50%, the recall is less than 90%. Therefore, in future works, we should further optimize the method to improve the recall of malicious nodes and reduce the FPR and FNR. For actual environmental monitoring, the possibility of misjudgment caused by malicious attacks is overall minimized.

Algorithm 1

Algorithm 1 – Introduction of DDF-2($\hat{X}^+_{k-1}$, $\hat{S}^+_{k-1}$).

Input: $\hat{X}^+_{k-1}$: the prior estimation of the (k-1)-th state; $\hat{S}^+_{k-1}$: the associate covariance of prior estimation error of the (k-1)-th state.
Output: $\hat{X}^+_k$: the prior estimation of the (k)-th state; $\hat{S}^+_k$: the associated covariance of prior estimation error of the (k)-th state.

1 //compute the corresponding first- and second-order divided differences
$2S^{(1)}_{X\hat{X},k-1} = \frac{1}{2h}(F(\hat{X}^+_{k-1} + h\hat{S}^+_{X,k-1}, W_{k-1}) - F(\hat{X}^+_{k-1} - h\hat{S}^+_{X,k-1}, W_{k-1}))$;
$3\ S^{(1)}_{XW,k-1} = \frac{1}{2h}(F(\hat{X}^+_{k-1}, W_{k-1} + hS_W) - F(\hat{X}^+_{k-1}, W_{k-1} - hS_W))$;
$4\ S^{(2)}_{XW,k-1} =$
$\frac{\sqrt{h^2-1}}{2h^2}(F(\hat{X}^+_{k-1}, W_{k-1} + hS_W) + F(\hat{X}^+_{k-1}, W_{k-1} - hS_W) - 2F(\hat{X}^+_{k-1}, W_{k-1}))$;
$5\ S^{(2)}_{X\hat{X},k-1} =$
$\frac{\sqrt{h^2-1}}{2h^2}(F(\hat{X}^+_{k-1} + h\hat{S}^+_{X,k-1}, W_{k-1}) + F(\hat{X}^+_{k-1} - h\hat{S}^+_{X,k-1}, W_{k-1}) - 2F(\hat{X}^+_{k-1}, W_{k-1}))$;
6 //compute the$\hat{X}^-_k$ and $\hat{P}^-_{X,k}$
$7\ \hat{X}^-_k = \frac{h^2-1}{2h^2}F(\hat{X}^+_{k-1}) + \frac{1}{2h^2}(F(\hat{X}^+_{k-1} + h\hat{S}^+_{X,k-1}) + F(\hat{X}^+_{k-1} - h\hat{S}^+_{X,k-1}))$;
$8\hat{S}^-_{X,k} = [\hat{S}^{(1)}_{X\hat{X},k-1}\hat{S}^{(1)}_{XW,k-1}\hat{S}^{(2)}_{X\hat{X},k-1}\hat{S}^{(2)}_{XW,k-1}]$;
$9\ \hat{P}^-_{X,k} = \hat{S}^-_{X,k}(\hat{S}^-_{X,k})^T$;
10 //compare the Euclid distance, D_k, with the threshold, τ;
$11 D_k \leftarrow |Z_k - \hat{X}_k|$;
12 If $D_k > \tau$, then
13 judge false;
14 else
15 judge true;
16 end if
17 //compute the first- and second-order divided differences
$18S^{(1)}_{Z\hat{X},k} = \frac{1}{2h}(g(\hat{X}^-_k + h\hat{S}^-_{X,k}, V_k) - g(\hat{X}^-_k - h\hat{S}^-_{X,k}, V_k))$;
$19S^{(1)}_{ZV,k} = \frac{1}{2h}(g(\hat{X}^-_k, V_k + hS_V) - g(\hat{X}^-_k, V_k - hS_V))$;
$20S^{(2)}_{ZV,k} = \frac{\sqrt{h^2-1}}{2h^2}(g(\hat{X}^-_k, V_k + hS_V) + g(\hat{X}^-_k, V_k - hS_V) - 2g(\hat{X}^-_k, V))$;
$21\ S^{(2)}_{Z\hat{X}^-,k} = \frac{\sqrt{h^2-1}}{2h^2}(g(\hat{X}^-_k + h\hat{S}^-_{X,k}, V_k) + g(\hat{X}^-_k - h\hat{S}^-_{X,k}, V_k) - 2g(\hat{X}^-_k, V_k))$;
22 //compute $\hat{Z}^-_k$ and $P_{XZ,K}$;
$22\ \hat{Z}^-_k = \frac{h^2-1}{2h^2}g(\hat{X}^-_k) + \frac{1}{2h^2}(F(\hat{X}^-_k + h\hat{S}^-_{X,k}) + g(\hat{X}^-_k - h\hat{S}^-_{X,k}))$;
$23\ S_{ZZ,k} = [\hat{S}^{(1)}_{Z\hat{X}^-,k-1}\hat{S}^{(1)}_{ZV,k}\hat{S}^{(2)}_{Z\hat{X},k}\hat{S}^{(2)}_{ZV,k}]$;
$24\ P_{XZ,k} = \hat{S}^-_{X,k}(S^{(1)}_{Z\hat{X}^-,k})^T$;
25 // calculate the Kalman gain, K_k, as well as $\hat{X}^+_k$, $\hat{P}^+_k$;
$30\ K_k = P_{XZ,k}(S_{ZZ,k}(S_{ZZ,k})^T)^{-1}$;
$31\ \hat{X}^+_k = \hat{X}^-_k + K_k(Z_k - \hat{Z}^-_k)$;
$32\ \hat{P}^+_{X,k} = (\hat{S}^-_{X,k} - K_kS^{(1)}_{Z\hat{X},k})(\hat{S}^-_{X,k} - K_kS^{(1)}_{Z\hat{X}^-,k})^T + K_kS^{(1)}_{ZV,k}(K_kS^{(1)}_{ZV,k})^T$;
33 return $\hat{X}^+_k$,$\hat{S}^+_k$;

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by Beijing Municipal Natural Science Foundation (No. 19L2020), and National Key Research and Development Project(Key Technologies and Applications of Security and Trusted Industrial Control System, No. 2020YFB2009500).

REFERENCES

Bhuiyan MZA, Wu J. Collusion attack detection in networked systems. In: IEEE 14th Intl Conf Depend Autonom Sec Comut; 14th Intl Conf Pervas Intell Comput; 2nd Intl Conf Big Data Intell Comput Cyber Sci Technol Congr; 2016. p. 286–93.

Gao L, Battistelli LG, Chisci L. Distributed joint sensor registration and multitarget tracking via sensor network. Inf. Fusion 2019;46(3):218–30.

Hammamouche A, Omar M, Djebari N, Tari A. Lightweight reputation-based approach against simple and cooperative black-hole attacks for MANET. Inf. Secur. Techn. Rep. 2018;43:12–20.

Hua Y, Chen F, Deng S, Duan S, Wang L. Secure distributed estimation against false data injection attack. Inf. Sci. 2020;515:248–62.

Jaint B, Singh V, Tanwar LK, Indu S, Pandey N. An efficient weighted trust method for malicious node detection in clustered wireless sensor networks. In: 2nd IEEE ICPEICES; 2018. p. 1183–7.

Kalkha H, Satori H, Satori K. Preventing Black Hole Attack in Wireless Sensor Network Using HMM. In: 2nd ICDS; 2018. p. 552–61.

Karthigadevi K, Balamurali S, Venkatesulu M. Based on Neighbor Density Estimation Technique to Improve the Quality of Service and to Detect and Prevent the Sinkhole Attack in Wireless Sensor Network. In: IEEE INCOS; 2019. p. 1–4. doi:10.1109/INCOS45849.2019.8951406.

Kumar B, Bhuyan B. Game theoretical defense mechanism against reputation based sybil attacks. Procedia Comp. Sci. 2020;167:2465–77.

Kumar M, Mukherjee P, Verma K, Verma S, Rawat DB. Improved deep convolutional neural network based malicious node detection and energy-efficient data transmission in wireless sensor networks. IEEE Trans. Netw. Sci. Eng. 2021. doi:10.1109/TNSE.2021.3098011.

Li L, Yang H, Xia Y, Yang H. Event-based distributed state estimation for linear systems under unknown input and false data injection attack - ScienceDirect. Signal Proc. 2020;170.

Liu Y, Ma M, Liu X, Xiong NN, Liu A, Zhu Y. Design and Analysis of Probing Route to Defense Sink-hole Attacks for Internet of Things Security. IEEE Trans. Netw. Sci. Eng. 2018;7(1):356–72.

Liu Y, Wang T, Zhang S, Liu X, Liu X. Artificial intelligence aware and security-enhanced traceback technique in mobile edge computing. Comput. Commun. 2020;161:375–86.

Martin VDG, Polinder H, Ferreira JA. Analysis and neutral voltage-based detection of inter-turn faults in high-speed permanent-magnet machines with parallel strands. IEEE Trans. Ind. Electron. 2015;62(6):3862–73.

Meng W, Luo X, Li W, Li Y. Design and evaluation of advanced collusion attacks on collaborative intrusion detection networks in practice. In: IEEE Trustcom/BigDataSE/ISPA; 2016. p. 1061–8.

Nobahary S, Babaie S. A credit-based method to selfish node detection in mobile ad-hoc network. Appl. Comput. Syst. 2018;23:118–27. doi:10.2478/acss-2018-0015.

Nørgaard M, Poulsen NK, Ravn O. New developments in state estimation for nonlinear systems. Automatica 2000;36(11):1627–38.

Spachos P, Hatzinakos D. Real-time indoor carbon dioxide monitoring through cognitive wireless sensor networks. IEEE Sensors 2015;16(2):506–14.

Tamandani YK, Bokhari MU. SEPFL routing protocol based on fuzzy logic control to extend the lifetime and throughput of the wireless sensor network. Wirel. Netw. 2016;22(1):647–53.

Tufail S, Batool S, Sarwat AI. False Data Injection Impact Analysis In AI-Based Smart Grid. In: IEEE SoutheastCon; 2021. p. 01–7.

Wang Y, Song Y, Lewis FL. Robust adaptive fault-tolerant control of multiagent systems with uncertain nonidentical dynamics and undetectable actuation failures. IEEE Trans. Ind. Electron. 2015;62(6):3978–88.

Xie L, Mo Y, Sinopoli B. False data injection attacks in electricity markets. In: 1st IEEE Int Conf Smart Grid Commun; 2010. p. 226–31.

Yao L, Wang DH, Liang X, Wan J. Research on multi-level fuzzy trust model for wireless sensor networks. Chi J. Sci. Instrum. 2014;35(7):1606–13.

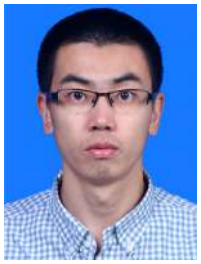
Yaseen Q, Aldwairi M, Jararweh Y. Collusion attacks mitigation in Internet of Things: a fog based model. Multimed. Tools Appl. 2018;77:18249–68.

Yingxu Lai received her Ph.D. from Chinese Academy of Sciences in 2003. She joined the College of Computer Science, Beijing University of Technology in 2003 and is current a full professor. She was a visiting scholar at Arizona State University from 2013 to 2014. Her research interest covers cloud computing, network security, edge computing and trusted computing.

She has had over 70 papers published in various international journals and conferences. She is currently an Associate-Editor of the *Journal of Artificial Intelligence and Technology*.

Liyao Tong received the B.S. degree from Inner Mongolia University of technology in 2018. She is currently working toward the M.S. degree with the College of Computer, Beijing University of Technology, China. Her research direction is malicious node detection in wireless sensor networks.

Jing Liu received her Ph.D. from Beijing University of Technology in 2017. She joined the faculty at Beijing University of Technology in 2002 and is currently a lecturer. Her research interests include information network security and trusted computing.



Yipeng Wang received the Ph.D. degree in computer science from Institute of Computing Technology, Chinese Academy of Sciences (CAS) in 2014. He is currently an Associate Professor with Faculty of Information Technology, Beijing University of Technology. His-research interests are in networking security and machine learning, in particular network protocol inference.

Tong Tang studied for a bachelor's degree from Beijing University of Technology in 2017. His-research direction includes Protection of Sensitive Data in Industrial Internet and malicious node detection in wireless sensor networks.

Zijian Zhao studied for a bachelor's degree from Beijing University of Technology in 2017. At present, his research direction includes Protection of Sensitive Data in Industrial Internet and malicious node detection in wireless sensor networks.

Qin Hua received the PH.D degree in computer science from Beijing University of Technology in 2007. She is currently a professor of BJUT. Her research interest includes applications of Telecommunication and Computer Networks, Internet of Things, Wireless Networks.