Review Article

Trust-Based Attack and Defense in Wireless Sensor Networks: A Survey

Weidong Fang, Wuxiong Zhang, Wei Chen, Tao Pan, Yepeng Ni, and Yinxuan Yang

1Science and Technology on Microsystem Laboratory, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 201800, China
2University of Chinese Academy of Sciences, Beijing 100049, China
3Shanghai Research Center for Wireless Communication, Shanghai 201899, China
4School of Mechanical Electronic & Information Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China
5School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China
6Shenhua Information Technology Co., Ltd, Beijing 100011, China
7Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou 350121, China
8School of Data Science and Media Intelligence, Communication University of China, Beijing 100024, China
9Fuzhou Internet of Things Open Lab Co., Ltd, Fuzhou 350015, China

Correspondence should be addressed to Wuxiong Zhang; wuxiong.zhang@mail.sim.ac.cn

Received 29 November 2019; Revised 2 August 2020; Accepted 18 August 2020; Published 10 September 2020

Copyright © 2020 Weidong Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a key component of the information sensing and aggregating for big data, cloud computing, and Internet of Things (IoT), the information security in wireless sensor network (WSN) is critical. Due to constrained resources of sensor node, WSN is becoming a vulnerable target to many security attacks. Compared to external attacks, it is more difficult to defend against internal attacks. The former can be defended by using encryption and authentication schemes. However, this is invalid for the latter, which can obtain all keys of the network. The studies have proved that the trust management technology is one of effective approaches for detecting and defending against internal attacks. Hence, it is necessary to investigate and review the attack and defense with trust management. In this paper, the state-of-the-art trust management schemes are deeply investigated for WSN. Moreover, their advantages and disadvantages are symmetrically compared and analyzed in defending against internal attacks. The future directions of trust management are further provided. Finally, the conclusions and prospects are given.

1. Introduction

Currently, the standardization works for Narrowband Internet of Things (NB-IoT) have been completed; the wireless sensor network (WSN) is taken as an important component for sensing and aggregating information. As the tentacles of social networks [1–4], the WSNs, which provide sensed information in context-aware and personalized social applications, have been widely deployed in many fields, such as smart cities, intelligent transportation, intelligent connected vehicles, precision agriculture, and environmental monitoring. Meanwhile, there are many research hotspots, including routing and access protocols, image recognition and target tracking, trusted transmission and trust management scheme [5], and energy consumption balance and energy efficiency. However, the information security is of mutual concern. In this regard, scholars focus on ensuring that sensed data is transmitted by the effective security schemes (e.g., secure routing protocol [6], security data fusion [7], and secure network coding [8]), to deliver to the end user in secure. The requirements for social network are shown in literatures [9–12]; the tasks and functions of WSN can be performed.
accurately and in real time, even though the network is being attacked by adversary.

Currently, the information security in WSNs is facing the enormous challenge, which comes from the security attacks including external attacks and internal attacks. The traditional security schemes (e.g., encryption [13] and authentication [14]) can only defend against the external attacks instead of the internal attacks. There are a few of studies demonstrate the trust management scheme is one of effective approaches to detect and defend against the internal attacks [15].

Trust management originated from sociology. In WSNs, in order to establish a secure communication link, it is necessary to guarantee that the intermediate nodes forwarding data packets are trusted in the network. Hence, it is essential to establish an effective trust model. In a trust model, each sensor node is allowed to evaluate the trustworthiness of neighbor nodes by interaction between nodes. Moreover, based on trust model, a trust management system is constructed to mitigate or defend against internal attacks, which are launched by captured or compromised nodes. In addition, trust management schemes are also used to evaluate the quality of received information, provide network security services including access control and malicious node detection, and secure resource sharing.

The research on trust management technology in WSNs is a challenging direction. How to construct a trust model is a key issue. By investigating and analyzing a large quantity of related literatures, these scholars mainly focus on two aspects: one is how to detect and defend against internal attacks, and the other is to obtain the trustworthiness of neighbor node to make decisions (e.g., selecting the next hop in secure and achieving the secure aggregation). Compared with the latter, we argue the former is more important.

Considering the above requirements and facilitating the research on attack and defense in the near future, in this paper, the trust management system and the typical internal attacks in WSNs are overviewed and investigated in Section 2. Furthermore, the state-of-the-art trust management schemes and trust models are deeply surveyed in Section 3. The detection and defense against security attacks with trust are comprehensively compared and analyzed in Section 4. Then, some valuable future research directions for trust management in WSNs are suggested in Section 5. Finally, the conclusions are drawn in Section 6.

2. Trust Management System and Internal Attack

In this section, the trust management system (TMS) is overviewed, and the internal attacks in WSNs are investigated.

2.1. Trust Management System. In general, there are five interrelated components in trust management system, including collecting, storing, modelling, transferring, and decision-making.

2.1.1. Collecting. It refers to collecting the trust elements, which involve the status of nodes’ interaction, location information, and sensed data. The reputation of the nodes is evaluated based on these collected trust elements. The trust value is further calculated from them. Therefore, the trust value becomes more accurately with more sufficient collected trust elements.

2.1.2. Storing. It refers storing trust element, trust values, and reputation. The storage must be systematically considered due to constrained resources for sensor nodes. Firstly, memory spaces would be impacted by the storage type of the sensed data. For instance, a float number consumes more memory than an integer. Secondly, the storage time of information would be considered; those outdated information should be emptied in time to save space. Finally, the location used to store information would be also concerned. In a clustered WSN, the trust value can be stored in the cluster head. When a cluster member needs to use the trust value, the cluster head may transfer it to this member.

2.1.3. Modelling. It refers to modelling the trust and reputation in WSNs, which is the key component of TMS. How to model needs to consider many factors, including the aging of trust value, whether to use indirect information, the weight of indirect information, the weight of each trust element, and the countermeasures aimed at defending against different attacks. In addition, the computational capabilities and energy supply of sensor nodes, and different network topologies must also be considered. Generally, the reputation model is a probabilistic statistical model, which is typical based on the beta distribution, the Gaussian distribution, or the binomial distribution.

2.1.4. Transferring. It involves reputation transfer and trust transfer between two nodes. The reputation transfer usually refers to when a node i need to evaluate the reputation of a node j, it initiates the reputation request to these common nodes (m1, m2, m3, and m4) between nodes i and j, and then they provide the reputation response of node j to node i. The process of reputation transfer is shown in Figure 1. The trust transfer is the Certificate Authority (CA) of the network provides the third-party trust value to the node, in order to complete the trust evaluation. For a hierarchical WSN, the CA is the cluster head, and the Base Station (BS) is CA in planar WSNs.

2.1.5. Decision-Making. Based on calculated trust value, the trust decisions should be made. Currently, decision-making with trust is divided into two categories as follows: (1) defending against the internal attacks: this is to punish a node with a low trust value. It is to directly drag it into the blacklist to exclude the network forever or make the node regain the trust based on the consideration of the selfish node and the energy consumption and (2) selecting the next hop in secure: in short, the trade-off between the security and performance should be comprehensively considered for the resource-constrained sensor nodes.

2.2. Typical Internal Attacks. The internal attacks are launched by the compromised or captured nodes. The attack behaviors involve discarding, replaying, tampering, and forging data packets, as well as providing the fake routing.
information. Since these malicious nodes have obtained the transmission schemes and held the key of the network, the internal attacks are more dangerous, and traditional encryption and other security mechanisms have no effect.

The typical internal attacks in WSNs are investigated and presented as follows: denial of service attack (DoS attack) [16], bad-mouthing attack [17]/slander attack [18], on-off attack [19], garnished attack [20], reputation time-varying attack [21], sleeper attack [22], conflicting behavior attack [23], Sybil attack [24], node replication attack [25], selfish attack [26], flooding attack [27], selective forwarding attack [28], black hole attack [29], ballot stuffing attack [30], collusion attack [31], sinkhole attack [32], data forgery attack [33], etc.

In the next sections, current researches on trust management scheme/trust model will be reviewed, and the capabilities defending against internal attacks with trust will be compared and analyzed.

3. Related Works

Currently, the research on the trust management mainly focuses on several aspects containing trust model, trust management scheme, and protocol optimization in WSNs (shown in Figure 2).

3.1. Trust Model. The trust model provides a framework for establishing and managing trust relationships between two nodes and ensures that the legal nodes can be trusted to participate in the process of information transmission.

Ganeriwal and Srivastava proposed a framework, which was based on RFSN (Reputation-based Framework for high integrity Sensor Networks) [34]. The framework consisted of five components including direct reputation evaluation, indirect reputation evaluation, reputation synthesis, reputation transfer and nodes’ behavior trust. Two important units in this framework were watchdog and reputation systems. The watchdog was used to monitor the behaviors of the neighbor nodes, especially to detect invalid information generated by abnormal nodes. It further classified these behaviors into cooperative or noncooperative behaviors. The reputation system was responsible for maintaining, managing, and updating the nodes’ reputation, in order to calculate the trust value. The reputation was generated by the observation of watchdog or integration according to other available information. For obtaining more objective trust value, the historical behaviors $R_i$ of sensor nodes were considered to calculate the current trust value. Therefore, based on a given reputation (node $i$ to node $j$), the trust value $T_{i,j}$ can be generated as follows:

$$T_{i,j} = E[R_{ij}] = E[\beta(E[\alpha_i, \beta_j])] = \frac{\alpha_i}{\alpha_i + \beta_j},$$

where $\alpha_i$ and $\beta_j$ represented the cooperative and the noncooperative numbers of node $j$ for node $i$. If the trust value was lower than a set threshold value, the node $j$ would be taken as abnormal, otherwise normal. RFSN provided a scalable scheme to detect the abnormal behaviors caused by malicious and erroneous nodes. Moreover, by introducing the aging factor, the historical behaviors were taken into the trust evaluation. Furthermore, based on RFSN, they also proposed a Beta Reputation System for Sensor Networks (BRSN) by using Bayesian networks. In BRSN, the feasibility of the beta distribution of node reputation was verified in the derivation process, and the calculation of reputation updating, aging, indirect information, and trust value and the updating and sintering the reputation were provided in detail. However, although the positive reputation information in RFSN was only transferred to mitigate the risk attacked by malicious nodes, the efficiency of the system was influenced inevitably. In addition, RFSN could not support the mobility of the nodes, and BRSN could not defend against the internal attacks with a high-reputation malicious nodes.

Yang et al. analyzed the impact on high-reputation malicious nodes and proposed a Multiple Attacks & Three Party-BRSN model (MA&TP-BRSN) [35] to improve BRSN. The proposed model was constructed by two components: one is MA-BRSN trust value calculation approach to solve the single detecting and evaluating attack issue in the existing reputation systems to a certain extent, and the other was TP-BRSN, which made the updating calculation of the third-party indirect reputation more objective, in order to achieve the defense against the internal attacks of the high-reputation malicious nodes. Yin et al. proposed an Improved BRSN (IBRSN) [36] for identifying the malicious recommendation and defending against the slander attack of high-reputation nodes. They introduced the indirect reputation of third-party nodes into IBRSN to eliminate the defects in BRSN to a certain extent. Jiang et al. proposed an Effective Distributed Trust Model (EDTM) for WSN [37]. The EDM was composed of three parts as follows: direct trust, recommended trust and indirect trust, and the direct trust and the recommendation trust were calculated selectively according to the number of received packets. In EDTM, when calculating direct trust, the communication trust, energy trust, and data trust were considered simultaneously and the
trustworthiness and familiarity were defined to improve the accuracy of the recommendation trust simultaneously. The trust value was calculated more comprehensively, the reputation of the sensor node was evaluated more accurately, and the malicious nodes were effectively prevented from destroying the network security in this model. However, the weights of various trusts needed to be further researched, and the threshold selection was a challenge.

In addition, some scholars also improved BRSN. Zhang et al. introduced the analysis of the social network relevance into the trust model based on BRSN and proposed the Sensor Node Trust Update Algorithm (SNTUA) [38] by using the “social network relevance”. In SNTUA, the reputations of nodes and their neighbors were further modified and comprehensively evaluated to improve the detection rate of malicious behaviors and reduce the impact of malicious nodes on WSN. Zhou and Shao proposed an improved trust evaluation model for WSN referred to as ZHOU [39] after the analysis of BRSN, based on Bayesian and entropy. In ZHOU, they considered the abnormal behavior brought by nonintrusive factors and introduced anomalous attenuation factor. Moreover, they used the modified Bayesian equation to estimate direct trust, updated it with sliding window and adaptive forgetting factor, and determined whether it was sufficiently reliable according to the level of direct trust as comprehensive trust. In this model, network energy consumption and the impact of malicious feedback were reduced. If a direct trust was not sufficiently trusted, indirect trust was calculated to obtain comprehensive trust. The entropy was used to assign weights to different recommendations. It could overcome the limitations brought about by subjectively deployed weights, in order to enhance the adaptability of the model simultaneously.

Chen et al. proposed an Agent-based Trust model for Sensor Network (ATSN) [40]. In ATSN, an agent node used the promiscuous mode to observe the behaviors of sensor nodes, which were divided into good behaviors and bad behaviors. Furthermore, the agent node calculated all good behaviors which were represented as $p$, and bad behaviors which were represented as $n$ separately; the reputation space was defined as follows:

$$RS = \{ (p,n) \mid p,n \geq 0 ; t = p+n \}.$$  

(2)

The trust domain was defined as $TS = \{ (pt,nt,ut) \}$, where $pt$, $nt$, and $ut$ represented the positive trust, negative trust, and uncertainty, respectively. In ATSN, the storage space and computational complexity could be minimized for the common sensor nodes. The trust value of the nodes was calculated to mitigate the slander attacks and on-off attacks by using the reputation of the direct neighbor node. However, the behaviors of neighboring nodes were difficult to be completely recorded, due to the data packet loss caused by the frequent communication or the hardware failure of cheap nodes. Hence, this would cause the trust and reputation system uncertainty. The security of ATSN relied heavily on the agent node, and the assumption that the agent node could defend against any security threat had no practical meaning. In addition, ATSN did not solve the issues of the updating trust and reputation.

Sinha and Jagannatham proposed a Gaussian-based trust and reputation management system for fading MIMO (Multiple-Input Multiple-Output) WSN [41]. Based on multivariable Gaussian distribution and Bayes’ theorem, the system considered the impact on MIMO wireless fading channels. Combining with direct and indirect reputation information, the reputation and trust value in this system were calculated, in order to effectively isolate malicious nodes. However, the calculation process was too complex to be suitable for resource-constrained sensor nodes. In addition, Zhang et al. proposed a dynamic trust establishment and management framework for clustered WSN [42]. They considered and introduced some new impact factors (such as nodes only communication with cluster head and using only the used cluster head reputation), by which made the system more secure.

Chen proposed a Task-based Trust framework for Sensor Networks (TTSN) [43], in which sensor nodes held the reputation of neighbor nodes with several different tasks to evaluate their trust. In TTSN, the trust was established by the task and trust management module, which consisted of three units: monitoring unit, reputation processing unit, and task and trust processing unit. The calculating trust approach referred to RFSN, and each sensor node had several trust values. Relatively speaking, TTSN was more suitable for the applications of large-scale WSN.

Zhu et al. proposed a Rank-based Application-driven Resilient Reputation framework Model (RARRM) in WSNs [44]. In RARRM, based on the driving of application program, the different ranks of trust values depended on different requirements.
Feng et al. proposed the Node Behavioral strategies Banding belief theory of the Trust Evaluation algorithm (NBBTE) [45], which was based on the behavioral strategy binding D-S (Dempster-Shafer) evidence theory. In this model, the sensed area with constrained resource was divided into few logical grids, and each grid was categorized with a unique identification. Then, the sensor nodes deployed in each grid verified location information of their neighbor nodes by using ECHO protocol. Each node further cross-checked the redundant sensed information of neighbors, and evaluated the trustworthiness of neighbors to detect the inconsistent data from malicious nodes. Finally, in the sink node, the sensed data from their grids could be aggregated and transmitted, and the inconsistent data from malicious nodes could be excluded simultaneously.

Hur et al. proposed a trust evaluation model to distinguish forged data of illegal nodes, so named DFDI [46]. In this model, the sensed domain with constrained resource was divided into few logical grids, and each grid was categorized with a unique identification. Furthermore, the ECHO protocol was used to verify the location information of the neighbors by deploying the sensor nodes in each grid. The sensor nodes cross-checked the redundant sensed information of neighbors and evaluated the reputation of neighbors based on their own checked results. The trust value was obtained by a weighted summation of the following three parameters: the consistency of the sensed information, the capability of communication, and the remaining duration of the node. At the sink node, the inconsistent data from malicious or compromised nodes could be detected by the transmitted aggregation result of each grid.

Fang et al. researched and found on-off attacks had greater concealment and aggression. Due to dynamically adjusting the reputation value, this attack was difficult to detect. Hence, a trust model based on a beta distribution that could defend against this attack was proposed (abbreviation: FANG) [47]. The different decision approach was adopted under the beta distribution. When the change of the trust value exceeded the set threshold, it indicated that the compromised node was launching the on-off attack. The scheme was easy to implement on resource-constrained sensor nodes. In addition, considering that the behaviors of the reputation-time-varying attack were similar to the impact of the mobile obstacle on the wireless signal transmission, they proposed a Time-window-based Resilient Trust Management Scheme (TRTMS) [21]. They further analyzed the behavior of normal/nodes and compromised nodes over a certain time interval and identified the abnormal trust values by the trend analysis. Simultaneously, they introduced control factors and time windows to detect and remove the compromised node that launched the reputation-time-varying attack from the suspected malicious nodes. The decision-making process is shown in Equations (3) and (4).

\[
N_R = \begin{cases} 
N_R + 1, & \text{if } \Delta T_{n-1} > 0 \text{ and } \Delta T_n \leq 0 \text{ and } \Delta T_{n+1} < 0, \\
N_R, & \text{else,}
\end{cases}
\]

\[
T_c = \begin{cases} 
T(n), & N_R < \tau, \\
0, & N_R \geq \tau,
\end{cases}
\]

where \(N_R\) was the reversed number of the trust difference and \(\Delta T\) was the change trend of trust. The misjudgments caused by the moving obstacles were solved by the TRTMS scheme effectively.

Xiao et al. researched the problem of Determining Faulty Readings (DFR) [48] and argued that arbitrary and noisy readings were fault readings. Furthermore, based on network correlation, they constructed the similarity between two sensor readings by exploring the correlation of sensor readings and then modelled it into a graph \(G = (V, E)\), where \(V\) represented the sensor network, and \(E\) represented the correlation between two nodes. If two neighbor nodes did not have any similarity in readings, then the two nodes were not directly connected. Once a similarity of the network was established, it was easy to infer the similarity between two sensor nodes. In addition, a correlation-based sensor rating scheme could be established by exploring the Markov chain in the network, where the sensor rating represented the reputation of the sensor node. They also proposed an effective intranetwork voting algorithm with trust to detect the fault readings based on sensor ratings. Although simply filtering and discarding abnormal readings might reduce the monitoring accuracy of the important events, it could be effectively avoided when using sensor rating scheme to detect the fault readings.

Inconsistent, unusual, or erroneous readings were usually caused by two different reasons, which include intentional misconduct and unintentional error. The former was mostly caused by malicious nodes, and the latter was caused by hardware failure or interference. The DFR-based approach focused on detecting the fault readings instead of processing them. In order to evaluate the reputation of sensor data properly, Gomez et al. proposed a new Mechanisms based on Data Life Cycle (MDLC) [49], which had three sensor data states: (1) unprocessed, (2) routed, and (3) processed. The data was sensed by the node without any additional routing or processing, and it was considered unprocessed. When the sensor data was transmitted to another node, it was taken as routed. Processed state referred to the fact that the sensed data was filtered, converged, or aggregated. In the mechanism, the trusts of unprocessed, routed, and processed data were calculated based on subjective logic.

Since the establishment processes of most trust models were only based on the interaction of neighbor nodes, this required a very important premise, that is, the sensed data was normal and the energy was evenly consumed. Once the sensed data and energy had a trust risk, a malicious, selfish, or low-competitive node that appeared in a WSN would result in a trusted node that was no longer trusted. To address the issue, Xiao et al. proposed a Trust Model based on Communication trust, Energy trust and Data trust (TMCED) [48] model. In this model, communication trust referred to the relationship value calculated by two cooperative nodes, and this calculation was derived from the successful interaction ratio. Energy trust referred to the remaining energy of a node, whether or not it was
sufficient to complete new communication and data processing tasks. It was calculated by

\[
T_d = W_1 T_f + W_2 T_n + W_3 T_v,
\]

where \(T_f, T_n,\) and \(T_v\) were the node fault-tolerant trust value, the trust value of event report, and the data consistency trust value, respectively. By using energy trust, TMCDE could effectively detect the DoS attacks. Once a malicious node launched a DoS attack, it would consume more energy, and the energy trust became lower than normal nodes. Hence, malicious nodes with lower energy trust would be more easily detected.

Nie proposed a Trust model of Dynamic optimization based on entropy method (Trust-Do) [50], which used Entropy theory to determine the node weights in each group. The standard deviation of Group Local Evaluation (GLE) was then calculated to reflect the overall expectations of all nodes in the group, as well as the Standard Deviation of Local Evaluation (SDL).

\[
\text{GLE}(t, \rho) = \frac{1}{m} \sum_{k=1}^{m} R_k(t) \times W_{\rho}(t),
\]

\[
\text{SDL}(t, \rho) = \sqrt[2]{\rho} \sum_{k=1}^{N} \left( z^2_{ik}(t) - z^2_{jk}(t) \right), \quad i \neq j,
\]

where \(W_{\rho}(t)\) was the weight vector under the group \(\rho\) and was determined according to the entropy size corresponding to the trust matrix element \(Z_{ij}\) of a group. Comparing SDL with GLE of a node, if a SDL was larger than GLE of the group to be \(\delta (0 \leq \delta \leq 0.5)\) times, it was divided into higher trust value packets; on the contrary, if SDL value was lower than the trust value of each node in a group, and the node was considered a malicious node. Although the model improved the detection capability of abnormal nodes, it did not consider the energy consumption.

Wu and Li established a Multi-domain Trust Management Model (MdTMM) [51] by using the classical interaction number as a mathematical model. This model was usually applied to a hierarchical RFID (Radio Frequency Identification) system. In the model, each RFID reader was taken as a sensor node, and each tag was equivalent to a data carrier. Each domain had a CA to authenticate the readers in its domain, monitor the current events, and detect the abnormal nodes. The D-S evidence theory and time windows were used to rank trust values, in order to effectively defend against information-based attacks including tampering attacks, replay attacks, and forgery attacks.

Gilbert et al. proposed a Time Series Trust Model (TSTM) [52] based on Toeplitz matrix and Trust based Auto Regressive (TAR) process, which was based on data prediction, and the effects of aggregation and reconstruction of Compressed Sensing (CS) were verified by various performance indicators and different attack models. Li et al. designed the Intrusion Sensitivity-based Trust Management Model (ISTMM) [53], which used machine learning technology to automatically assign intrusion sensitivity based on expert knowledge. The performance of three different supervised classifiers in assigning sensitivity values was compared during the evaluation process.

Considering the existing universal trust model was difficult to meet the requirements of multihop routing, Liu et al. proposed a Trust Model based on Bayes Theorem (TMBBT) [54] for the multiple paths in WSNs. In this model, all nodes were divided into two categories: one was the nodes communicated with other nodes only via one-hop routing; the others were not only communicated via one-hop routing but also via multihop routing for one-hop unreachable. The trust evaluation consisted of two parts: communication trust and data trust. Communication trust was calculated based on cooperative routing information. The reputation and trust of the data depended on the ratio of data successfully received. This was due to the fact there were only direct communication and data instead of indirect communication and data; it could reduce the energy consumption. However, the calculations of trust value were not accurate enough without neighbors’ recommendations. In addition, how to combine communication trust with data trust was not mentioned in this article.

Zhang et al. proposed a novel scheme to detect the malicious node based on DPAM-MD (Density-based Partitioning Around Medoids-Malicious node Detection) algorithm [55]. In this scheme, a subaggressive node could be detected by combining Manhattan metrics and DPAM (Density-based Partitioning Around Medoids) algorithm on the basis of the traditional reputation threshold judgment model. Moreover, combining the intercluster with intracluster distance equalization objective functions, a novel density-based clustering algorithm was proposed to classify all nodes. It could effectively shorten the clustering time and improve the efficiency detected malicious node, especially for those obvious compromised nodes. Zeng et al. proposed a Gray Markov-based Model to improve BRSN (GMM-BRSN) [56] and then designed a query routing protocol to address the issue of Selective forwarding attack in the routing protocol on the basis of GMM-BRSN. The GMM-BRSN had higher security and lower energy consumption.

Atakli et al. proposed a Weighted-Trust Evaluation (WTE) scheme [57] for hierarchical WSN to detect malicious nodes. In WTE, the weighted trust was calculated as follows:

\[
W_n = \begin{cases} 
W_n - \theta \times r_n, & \text{if} \ U_n \neq E, \\
W_n, & \text{otherwise},
\end{cases}
\]

where \(U_n\) was the sensing data of the evaluated node, \(E\) was the aggregated data of cluster head, and \(\theta\) was the penalty ratio. \(r_n = m/s,\) where \(m\) was the number of nodes that produces inconsistent data and \(s\) was the total number of nodes under the cluster head. This scheme had higher security, when there were a small number of compromised nodes in the network; however, when more than a quarter of the nodes were compromised, the performance was unsatisfactory.

Mahmud et al. used an Adaptive Neural-Fuzzy Inference System (ANFIS) and brain-inspired trust management
model (TMM) to enhance the security of IoT devices and relay nodes [58]. The TMM could detect the malicious nodes in the network and utilize both node behavioral trust and data trust to evaluate the nodes trustworthiness. Chen et al. proposed [59] a trust evaluation model, which directs data trust compared real-time monitoring data with historical data. If the value was large, it was considered an abnormal node.

Karthik and Ananthanarayana [60] focused on data trust model, which was called as KARTHIK, especially data fault detection, reconstruction, and quality estimation for reliable event detection involved with Temporal, Spatial, and Attribute data modelling. The correlation of data in multiple dimensions including time and space was calculated to find faulty data. In terms of data trust, the calculation of coefficients was mapped to three integers of -1, 0, and +1, which represented data errors, uncertainty, and complete trust. Liu and Cheng proposed [61] a state space modelling approach for trust evaluation that employs a state space model for time series analysis. This model was named LIUCHENG by us. The trustworthiness of each node was modelled by a trust index; under the state, it formed a vector. Then, based on improved particle filter, the high-dimensional spatial trust value was calculated to better detect erroneous data. A certain amount of storage spaces and computational capabilities were required both in time and space. Singh and Verma presented a trust model for Flying Ad hoc networks (FANET). We called this model as KULDEEP [62], which consisted of QoS (Quality of Service) trust and social trust, which synthesized trust values through fuzzy logical classification and weight assignment. The features of node contained signal strength, packet delivery ratio, node's energy, and transmission delay, were calculated by percentage. The trust value calculation of the model involved all aspects of transmission to the path consumption and could provide protection against most internal attacks while ensuring network load balancing.

Ghugar et al. proposed a protocol Layer trust-Based Intrusion Detection System (LB-IDS) to secure WSN by detecting the attackers at different layers [63]. The trust value of a node was calculated by using the deviation of trust metrics at each layer with respect to the attacks. They also considered trustworthiness in PHY layer trust, media access control (MAC) layer trust, and network layer trust. Finally, the overall trust value of a node was estimated by combining the individual trust values of each layer. By applying the trust threshold, a sensor node was determined as trusted or malicious. The proposed system could defend against jamming attack at the physical layer, back-off manipulation attack at the MAC layer, and sinkhole attack at the network layer.

Zhao et al. proposed an Exponential-based Trust and Reputation Evaluation System (ETRES) to evaluate the trust and reputation of a node in WSNs [64]. ETRES was used to observe the nodes’ behavior, and exponential distribution was applied to represent the distribution of nodes’ trust. The trust of the node was used to look for reliable nodes to transmit data and weaken malicious attacks in WSNs. More significantly, the entropy theory was used to measure the uncertainty of direct trust values. Indirect trust was introduced to strengthen interaction information when the uncertainty of direct trust is enough high. In addition, the confidence factor was redefined, which could dynamically adjust the node trust value to weaken the harmful effects of the compromised nodes.

In ETRES, the exponential distribution was applied to represent the distribution of the reputation of a nodes, and the node’s behaviours were used to calculate the trust value, which involved the direct trust value and the indirect trust value. More significantly, the entropy theory was introduced to measure the uncertainty of direct trust values. The indirect trust value was adopted to strengthen the certainty of the trust value, when the uncertainty of direct trust was enough high. In addition, a confidence factor was redefined to dynamically adjust the trust value of a node, in order to weaken the harmful effects of the compromised nodes. The ETRES was used to look for secure relay nodes to forward data and prevent the malicious attacks in WSNs.

3.2. Trust Management System/Scheme. Since trust management scheme in WSNs was limited by hardware resources of sensor node, more behavior-based trust management schemes were adopted. These schemes were suitable for addressing the distributed authorization issues, and they had the advantages of flexibility and scalability.

Zhou et al. proposed a a trust and reputation management scheme for cluster-based WSN. [65]. In this scheme, the cluster head elected a node as a Surveillance Node (SN), which monitored the behaviors of cluster member nodes, calculated their reputation and trust, and evaluated their trustworthiness. The cluster head used this information to obtain the trust value of each node, in order to defend against attacks. In addition, a sensor node with higher trust value had a great opportunity to become a SN, thus enhancing the security of this cluster.

Boukerche et al. proposed an Agent-based Trust and Reputation Management scheme (ATRM) for wireless sensor nodes [66]. In ATRM, the trust and reputation were managed by minimized additional messages and time latency, and the trust and reputation information of the node were required to store as t-instrument and r-certificate. Since a node could not manage and calculate its own trust and reputation, each node was also required to have the ability to manage the trust and reputation of its host nodes. Moreover, any transaction was defined as an interaction between two nodes (requestor and provider). It was triggered by the requester, and then the provider chose to accept or reject. Before any interaction, the requester directly queried the local mobile agent to obtain the provider’s r-certificate. Depending on the provider’s certificate, the requester decided whether to start the interaction. When the interaction was complete, the requestor evaluated the provider’s trust based on QoS obtained in the interaction and sent the t-instrument to the provider’s local mobile agent. Based on the t-instrument collected, the mobile agent periodically released the r-certificate updated by its managed nodes. The advantage of ATRM was that there was no need to
centrally store trust and reputation, and the nodes provided their own reputation information when it needed. However, the establishment of ATRM required extraordinary assumptions. It assumed that the mobile agent was resilient to any threat, and the mobile agent was resilient to malicious nodes, which tried to steal or modify the information that the agent carried. The feasibility of these assumptions needs further research.

Yao et al. proposed a Parameterized and Localized trust management Scheme (PLUS) [67]. In PLUS, each sensor node held highly abstract parameters to evaluate the trustworthiness of the interested neighbor nodes, in order to detect the malicious nodes. Specifically, the direct trustworthiness of a node was calculated by the availability of the node and the proportion of the correct grouping. The indirect trustworthiness was calculated based on the neighboring signal value and the number of neighbors. The direct and indirect trustworthiness were synthesized according to different weights, in order to obtain the total trustworthiness. The PLUS was further used to design a routing scheme, named PLUS_R. In PLUS_R, all important control packets generated by the Base Station must contain a Hash Sequence Number (HSN), so that effectively guaranteed their integrity. However, the HSN increased the packet length and the energy consumption of transmission. Since the integrity of a packet was always checked, if checked fails, regardless of whether this packet was maliciously modified by the node, the trust value of this node would be reduced. Thereby, a normal node might be unfairly penalized.

Shaiikh et al. proposed a Group-Based Trust Management Scheme (GTMS) [68], which obtained a single trust value in the whole group. In GTMS, the trust value was calculated based on direct and indirect observations. The direct observations referred to successful and unsuccessful interactions, and indirect observations indicated the recommendations of trusted nodes with respect to particular nodes. The interaction referred to the cooperation of two nodes. When a node successfully received a packet, it would send back an ACK to the transmitter. If the transmitting node did not receive the ACK within a predefined threshold time, the data packet would be retransmitted. If the receiving node did not receive the retransmission of the packet within the threshold time of its neighbor node or found that the eavesdropping packet was illegally manufactured, the transmitting node would consider the interaction unsuccessful. If the number of unsuccessful interactions increased, the transmitting node reduced the trust value of the neighbor node and treated it as a malicious node. Compared with the traditional trust management scheme, GTMS focused on the trust value of a set of sensor nodes, rather than always focusing on the trust value of each node. GTMS not only provided a detection scheme for malicious nodes but also provided a certain degree of prevention scheme. Although GTMS took energy consumption into account, reduced the computing and communications expenditure of trust evaluation. However, it relied on a broadcast-based policy to collect many feedbacks, which in turn consumed additional resources and energy at another communication level.

He et al. proposed an attack-Resistant and lightweight Trust management scheme (ReTrust) [69]. In this scheme, a two-layer architecture was composed of the master node and sensor node, and the master node of each cell would manage the trust records of other master nodes and sensor nodes in this cell. Two network topologies were used, which involved an intracell topology and intercell topology. The former managed trust records for sensor nodes in this cell based on past direct interactions, and the latter managed the trust records of other master nodes through direct historical observations, recommendations, and indirect interactions. In addition, an aging parameter was also introduced, which assigned different aging factors to each historical moment in the evaluation window. ReTrust was lightweight and did not add any additional expenditure on resource-constrained sensor nodes; the trust calculation of the master node was simple. ReTrust could not only effectively identify malicious behaviors and eliminate malicious/fault nodes but also significantly improve network performance. However, the drawback of ReTrust was that the master node must have abundant storage resources and energy. Sensor nodes with limited resources did not have the ability to manage trust records of other nodes.

Yu et al. summarized Trust and Reputation Management (TRM) system in wireless communication systems [70]. They divided the existing TRM systems into two categories: the individual-level trust model and the system-level trust model. The individual-level trust model focused on the trust evaluation from one node to another. The system-level trust model included trust and reputation evaluation model and protocol. In TRM systems, by using an example of the individual-level trust model, they provided the trust and reputation of the initial phase, evaluated the reputation of the synthesized the direct and indirect reputation, and guided the trust evaluation and decision-making. In addition, the rewards and punishments in the system were based on the trustworthiness of nodes; several reward and punishment schemes for the system-level trust model were given. Duan et al. proposed an energy-aware trust derivation scheme with the game theory [71]. They analyzed the requirements of the network security and introduced the Trust Derivation Dilemma Game (TDDG) to design a risk model, in order to get the optimal number of collaboration nodes by encouraging the cooperation between nodes. The game theory was also used for trust derivation, which reduced the calculation cost. Li et al. proposed a Lightweight and Dependable Trust System (LDTS) for clustered WSN [20]. In LDTS, they proposed a lightweight trust decision-making scheme based on the node identity of a clustering WSN, to improve the system efficiency and reduce the harm of malicious nodes by eliminating the interactive feedback between cluster members and cluster heads. Since the cluster heads undertook many important tasks of data forwarding, they defined the trust evaluation method for the interaction between the cluster heads and the adaptive weighting approach. In addition, considering that the traditional entity based trust evaluation scheme was not suitable for the data-centric sensor network, Li et al. proposed a Data-centric Trust for Sensor Network (DTSN) scheme [72]. Simultaneously, a new approach,
Proof-of-Reputation-Relevance (PoRR), was presented to realize DTSN. Zia and Islam proposed a trust scheme based on Communal Reputation and Individual Trust (CRIT) [73]. In this scheme, the behavior of the nodes was monitored by watchdog, and each node held a trust and reputation table for evaluating its neighbors. Fang et al. proposed a multifactor reputation management scheme [74]. The multifactor involved event perception, packet forwarding, and data aggregation. The proposed scheme could be used to SPIN protocol to improve the data forwarding rate and delivery success rate in distrusted environment.

Fang et al. proposed a beta distribution-based Trust and Reputation Evaluation System (BTRES) for WSN [75] to address the security issue, which was vulnerable to be attacked from compromised nodes. Based on the interaction information between the nodes, in BTRES, the beta distribution was used to emulate the reputation of nodes, and the trust value was further calculated to obtain. In addition, weights and thresholds were used in combination to construct BTRES. The simulation results had shown that BTRES could effectively defend against the internal attacks and enhance the network security. The trust value of the node in BTRES could be used for the routing protocol or the aggregation scheme. When selecting routing or aggregating information, the node with the current high trust value was firstly selected, so as to ensure the security of information forwarding and transmission. Furthermore, they proposed a Binomial-Based Trust Management System (BTMS) [76] for WSN. The BTMS could only transfer the positive reputation between nodes, so as to mitigate the slander attacks.

Srinivasan et al. proposed a Distributed Reputation-based Beacon Trust System (DRBTS) [77] to detect and remove malicious beacon nodes provided incorrect location information. In DRBTS, the beacon nodes could be monitored each other, and the relevant information was provided for sensor node to select the competition trust. Every beacon node would monitor its neighbor nodes, observe them whether cheated, and update corresponding beacon node reputation in neighbor reputation list. After the error of indirect information of the beacon node was detected, the reputation of the neighbor node could be updated by using it. A sensor node deployed the neighbor node reputation list to decide whether used beacon position information based on simple majority vote scheme. In DRBTS, an undirected graph was built by using a network model, to synthesize the direct information and indirect information into the trust.

Karthik and Ananthanarayana proposed a Hybrid Trust Management Scheme (HTMS) for WSN [78]. In HTMS, it assumed that the network needed to evaluate the degree of trustworthiness of the nodes when it made decisions. Moreover, all trust score was obtained based on the trust component. Therefore, the data quality and transmission trust were considered. By detecting data errors with time-space correlation, the transmission trust and original data were used to estimate the trust score of the intermediate node and information trust score. And then the data trust score was used to make decisions. The direct trust was calculated based on the number of successful interactions. The data trust depended on whether the acquired sensory data was within the predictable scope, and mapped it as three integers: +1, -1, and 0. In addition, they also considered the residual energy level of the node and the uncertainty of the data. The correlation coefficient of the neighbor nodes was calculated by the association between node data in time-space and used as a positive correlation indicator for data trust. By using HTMS, some internal attacks including DoS attacks, bad-mouthing attacks, on-off attacks, attack on information, selective forwarding attacks, replication attacks, Sybil attacks, and collusion attacks were detected and defended against effectively. By setting a certain reward and punishment system, a reliable node and its source node were increased or decreased, and the trust score of the intermediate node could effectively detect those malicious, error, and selfish nodes.

Singh et al. propose a Light Weight Trust Scheme (LWTS) for clustered WSN [79]. In LWTS, each node would monitor the neighbor nodes. The monitoring events divided into two categories: success and failure. If the result of the monitoring event was a predictable result, then the event was taken as a successful interaction. Different from LDTS, the data package, control packet, and their message precision were included in trust measurement for LWTS. All calculation matrix dimensions were based on multiple neighbors of a node. Furthermore, they also considered the positive and negative feedbacks. It could defend against bad-mouthing attack to a certain extent and also consider the energy consumption of the node. However, most of the trust values were deduced based on the form of (S_{xy} + U_{xy})/(1 + S_{xy} + U_{xy}) as well as the traditional aging factor. It updated the trust value at different time. Although this scheme could defend against some internal attacks, including bad-mouthing attack and black hole attack in a certain extent, there was a lack of response speed to the attack.

Talbi et al. proposed an Adaptive and dual Data-Communication Trust scheme (ADCT) [80]. In a hierarchical network, a new communication trust T was defined according to the classic interaction number calculation equation:

\[
P = \frac{S(t)}{S(t) + U(t)},
\]

\[
T = \left[ 10 \times p^{(1-p)a} \right],
\]

where S(t), U(t), and P represented successful and unsuccessful communications in the time period (t), and the percentage of succeeding corresponding, respectively; \([\cdot]\) represented the integer function latest; \(a \geq 2\) was the parameter which affects the order of severity of trust function. The data trust feedback \(T_{ch}\) was built as follows:

\[
T_{ch} = \left[ 10 \times \frac{P_{\text{recommendations}} + 1}{P_{\text{recommendations}} + N_{\text{recommendations}} + 2} \right].
\]

\(P_{\text{recommendations}}\) and \(N_{\text{recommendations}}\) represented positive and negative data trust recommendations, respectively.
In ADCT, the duality data communication was used to deal with the unreliable recommendation, in order to establish the feedback from a cluster member to the cluster head. Therefore, it could prevent the recommendation of a harmful node and reduce the communication energy consumption. However, they made the decision without considering the dynamic cluster group (unset boundary) and united node energy level.

Reddy et al. used the D-S evidence theory to propose a communication and data trust for WSN (TWSN) [81]. In this scheme, the direct trust was set up on the number of forwarded packets \((p_i)\) and the number of packet loss \((q_i)\) in a certain moment of a node. Specifically, they compared the relevant Forwarding Ratio (FR) \(FR(t) = (p_i)/(p_i + q)\) in a certain moment with last moment, calculated the fluctuation of node forwarding consistency, and dealt with it by penalty factor or excitation factor. Based on the root mean square error, the similarity parameter was customized to correct the recommendation. Among them, the indirect trust weighted and summed multiple recommendations by using the evidence theory and the similarity parameters. Data trust was calculated based on the mean of the sensor data. Moreover, based on the comparison of the size of the sensor values, controlling data trust was increased and decreased by generating two factors after comparing the sizes. This scheme could be done without increasing the time window to realize better control effect of trust value. Combining the screening for recommendations, it could defend against the bad-mouthing attacks and on-off attacks to a certain extent.

Jin et al. proposed the Multi-agent trust-based intrusion detection scheme (Multi-agent) [82]. In this scheme, the data trust included four dimensions (packet loss rate, packet transmission frequency, packet receiver frequency, and energy consumption rate) and considered the speed of energy consumption. Therefore, a more energy-intensive attack such as DoS attacks or flood attacks could be detected by using this scheme.

Firoozi et al. [83] proposed a trust scheme in hierarchical networks, in which the cells were divided evenly by grid in plane space, and the data in a cell were processed. The cell distance and number of nonempty cells were defined for processing. And special situations were taken into consideration in CoSLIP, namely, an SL-(subjective logic-) based in-network data processing scheme for collocated WSNs. Combined with trust management, Janani and Manikandan proposed a secure PKI (Public Key Infrastructure) system [84] called as JANANI. By evaluating the hybrid trust value with the trust evaluation vector method, this scheme was effectively integrated into the hexagonal clusters to secure the PKI framework and detects and classified the misbehaviors, either selfishness or malicious, to take revocation actions on those nodes.

Meng et al. deployed a trust management application into [85] IoT in hierarchical networks; \(k\) paths were generated and the cuckoo search algorithm was used to find the optimal path. Combined with the Bayesian based on wireless traffic sampling, it could reduce the excessive data input of IoT devices to defend against black hole attacks and selective forwarding attacks. Sahoo et al. put forward a trust management focus on penalty and reward policy, named RASHMI [86]. Calculating the current time window to set dynamic parameters, RASHMI could defend against reputation time-varying attacks, especially on-off attacks. In RASHMI, the nodes were divided into benevolent/legitimate nodes, persistent malicious nodes, and transient malicious nodes. Then, the direct trust value was calculated using sliding time windows, fractions, and weighted summation root mean squares. Mathematically, size of dynamic timing sliding window was defined as ON period. As for reward and penalty schemes, \((S_{i,j}(t))(1 + S_{i,j}(t))\) signified the reward factor; \((U_{i,j}(t))^{-1/2}\) signified the punishment factor. Combined with the time window and reward penalty scheme, it could well control the trend of trust value and better detect and discover reputation time-varying attacks. The downside was that the recommendation for trust values and the fusion was weak, and the resistance to similar bad-mouthing attacks was weak.

Khan put the trust management scheme into practice in IoT, namely, called ZEESHAN [87]. In ZEESHAN, the beta distribution was used to calculate the trust value. Combined with the energy-limited IoT device, three different packet forwarding scheme algorithms were set to reduce the corresponding node energy consumption NLD (no listening for data forwarding), LDF (listen own data forwarding), or LT (listen to all transmissions).

Yang et al. put a novel application into Vehicular Networks [88] blend with blockchain. We called this scheme as YANG. This application inherited the decentralization and taper resistance of blockchain. All nodes or RSUs (RoadSide Units) collaboratively maintained an updated, reliable, and consistent trust blockchain, so that this system could resist message spoofing attacks, bad-mouthing attacks, and ballot stuffing attacks. Excessive computational capabilities, storage spaces, and energy resources were often required to send encrypted data and calculate hash values. Therefore, the application was limited to RSUs and deployed vehicle network scenarios with sufficient resources. Whether it was an internal attack or an external attack, the combination of blockchain and trust management made the trust management system more secure and reliable.

Smithamol and Rajeswari proposed a trust management middleware (TMM) [89], which applied in service selection in the cloud. The criteria of trust evaluation included CPU percentage, disk read throughput, disk write throughput, and network bandwidth, after the service filtering and selecting, and then through the OTA (Overall Trust Algorithm) with dynamic weight to calculate the overall trust value. This system could defend against the internal attacks including the QoS attacks and bad-mouthing attacks.

Pham and Yeo presented a trust management system that context-aware trust management scheme [90], which was named THI-CHAI. In this scheme, the nodes could be allowed to evaluate the trustworthiness of receiving events by considering the entity reputations of the senders under the vehicle networks. First, it utilized BF-based PSI to enable a node A to recognize the node B trust level. With a decision tree that estimates the entity trust adaptively to the available link ability information with encryption technology, which
means this system can resist the data-relevant attack, such as tampering attack.

For detecting on-off attack in health WSNs, Fang et al. proposed a Binomial Distribution-based Trust Management Scheme (BDTMS) [91]. Firstly, time interval between the highest trust value \( T_h(i) \) and the next highest trust value \( T_h(i+1) \) as a detection period \( P(i) \) was defined. There was the lowest trust value \( T_l(i) \) in a detection period, and this moment represents TIM. Secondly, then presented a descent time \( T_d(i) \), which was a time interval from \( T_h(i) \) to \( T_l(i) \), as well as an ascent time \( T_a(i) \) from \( T_l(i) \) to \( T_h(i+1) \). Finally, they gave any trust value \( T_a(i,m) \) during a descent time and any trust value \( T_d(i,n) \) during a descent time. If the following relationship was satisfied, the malicious node that launched the on-off attack can be basically detected. In Equation (10), \( F_d \) was the detection flag. If \( F_d \) was 0, the detected node was malicious; otherwise, it was a normal node. For malicious nodes, they would be remove from the routing table to achieve the defense against On-Off attack.

\[
F_d = \begin{cases} 
0, & \text{if } \left| T_d(i) - T_l(i) \right| < \delta, \\
\frac{T_d(i,m) - T_d(i)}{td(i)}. & \text{or } \sum_{k=0}^{T_a(i,TIM+k)} \left( T_a(i,TIM+k) - T_d(i,TIM-k) \right) < \sigma, \\
1, & \text{otherwise.}
\end{cases}
\]

In addition, Ukil proposed a collaborating computing model based on trust and reputation to detect and prevent the malicious attack [92]; this approach realized the choice of an optimal path, and enhanced the reliability. Ishmanov and Kim proposed a secure trust evaluation scheme to limit the increase in the trust value of malicious nodes for WSN [93]. Different from traditional trust management scheme, the proposed scheme was considered as the influence of abnormal node behaviors.

### 3.3. Protocol Optimization

Generally speaking, the protocol optimization referred to design the trust management protocol, in order to implement interaction with trust management related information. On the other hand, it referred to security optimization for routing protocols, transmission protocols, and data aggregation protocol by using the trust decision.

Bao et al. proposed a trust-based intrusion detection and Hierarchical Trust Management Protocol (HTMP) [94] in WSNs. The scheme was suitable for the routing protocol based on trust of intrusion detection. Furthermore, they analyzed the different influence on the choice of the minimum trust threshold value.

Gheorghe et al. proposed an Adaptive Trust Management Protocol (ATMP) [95], which was based on the behaviors of nodes to adjust the trust and reputation value. It included three phases: learning phase, exchange phase, and update phase. Learning phase got through the experience received from TinyAFD (Tiny Attack and Fault Detection framework) and judged the node’s behavior that was good or bad. Exchange phase was the empirical interaction between two neighbor nodes. Update phase was used to update the reputation and trust value with experience. The adaptivity of ATMP was from experience, and it adjusted reputation and trust value according to the behavior of the sensor node in each cycle. ATMP was interoperability, which embodied in proceeding exchange of respective behavior in exchange phase. Due to the adaptivity and interoperability of ATMP, it could defend against the internal attacks preferably. Tajeddine et al. propose CENtralized Trust-Based Efficient Routing Protocol (CENTER) [96]. CENTER took advantage of the information provided from BS, to detect and forbid the badness node which hampers or abuses network function. In CENTER, the BS collected the observed information of every node, and after several observations and calculations, a more accurate global network map was obtained. Furthermore, the BS estimated its service life on account of the condition of node activity, computing node behavior message (malicious, collaborate, compatibility), evaluating the trust value of every node (data trust and transmit trust), and took advantage of effective decision-making system to isolate malicious node of the network.

Priyoheswari et al. proposed a topology management route based on trust [97]. They used the Received Signal Strength Indicator (RSSI) as a characteristic parameter to join the calculation of trust value, in order to estimate the topology of WSN. This protocol could detect the behavior of abnormal node effectively. Mehetre et al. aimed at the internal attack of cluster WSN, used two-stage security scheme and dual assurance scheme, and proposed Trustable and Secure Routing Scheme (TSRS) [98]. Based on initiative trust, TSRS achieved to guarantee route protocol, to defend against a few internal attacks, such as black hole attacks and selective forwarding attacks. By using trust and cuckoo search algorithm to recognize trusted path, this scheme could combine energy selection and provide a secure route path. The scheme also offered the guarantee to prolong the network lifetime.

Chen et al. proposed the Peer-to-Peer (P2P) trust management protocol based on the elliptic curve [99]. This protocol provided the function of authentication and signature to protect the process of the trust value queries and rating reports. Furthermore, the protocol also generated two verified pseudonyms to take the place of node identity, of which one pseudonym was used for events and another pseudonym was used for the peer establishment procedure. Addo et al. proposed a Secure, Private and Trustworthy Protocol (SPTP) [100] to solve the issues of the security, privacy, and trust with mobile and cloud services in a Collective Intelligence (CI) scenario. Shilpa and Ambareesh proposed a trust management protocol in WSNs [101]. The protocol consisted of four parts: trust constitute, trust aggregation, trust formation, and application-level trust optimization design. It combined QoS trust with social trust to obtain a composite trust metric. In addition, the protocol allowed setting best trust in the trust aggregation process, to make subjective trust close to
objective trust in the individual’s trust attribute, and realized the minimum of trust deviation.

For trusted routing protocol, ahhal et al. proposed Trust-based Cross-Layer Model (TCLM) [102], which used the concept of cross layer (ACK from data link layer and TCP layer) to design a trust-based model for sensor networks, in order to isolate malicious nodes. Among them, data-packet statistics could be used to calculate values related to neighbor nodes, namely, trust value (denoted as $t$) and treatment ratio (denoted as $r$). The trust value characterized the degree of belief that neighboring nodes were reliable relative to packet delivery. The treatment ratio represented the statistical confidence in this trust. Let $L$ be the accumulation of packets forwarded by a sensor node and $N$ for the cumulative total of packets forwarded by the sensor node. Trust ($t$) and treatment ratio ($r$) are defined as follows:

$$t = \frac{L}{N},$$

$$r = 1 - \sqrt{\frac{12L(N-L)}{(N+1)N^2}}.$$  

Wang et al. created an Energy-efficient Trust Routing Mechanism named ETMRM [103]. They firstly extended the sensor flow tables to realize a lightweight trust monitoring and evaluation scheme at the node level, and detected and isolated the malicious nodes based on the trust information collected from sensor nodes. Under this message scheme, neighbor nodes’ report messages were aggregated and reported to reduce the size of the packets and the times of forwarding, so that to save energy and ensure the transmission of control traffic.

In addition, Gerriagaitia et al. proposed a reputation-based intrusion detection system for WSN [104], which analyzed and ensures the source of malicious attacks by using the trust values of different nodes. Ukil proposed a computational approach based on trust and reputation cooperation in WSNs [105], which effectively eliminated malicious nodes with high probability. They found secure forwarding paths among routes, so the approach had good trustworthiness and communication efficiency.

The trust theory originated from sociology. Generally, trust was considered a dependency. Interdependence meant that an interaction relationship existed in two parties. Regardless of the interacted content, it meant that the two parties have at least a certain degree of benefits, and their own benefits to be achieved must rely on the other party [106]. In a distributed network system, trust was defined as a subjective judgment of honesty, security, and trustworthiness, which made by an entity to another entity through observation and historical experience over a given period of time and context. Briefly, trust was a security scheme to defend against internal attacks and realize network self-healing. In WSNs, trust usually refers to predicting the credibility of future behavior of a node. The operation and acquisition of the trust value could only be obtained from sensing data directly, or the recommendation of the neighboring node, which generally changed with the behavior of the node. The trust value was usually used to determine whether the information was interacted between nodes. Moreover, the computational complexity of trust management in WSNs was related to many aspects, which involved different reputation distributions, node behavior trust/data trust, the coupling of direct trust and indirect trust of attack characteristics model, timeliness of trust information, and openness of wireless channels.

4. Security Analysis of Trust Management Technology

The research on the trust management is to detect malicious nodes and defend against internal attacks, in order to enhance the network security. For example, if a malicious node does not forward the received information, the trust value will decrease. The malicious node can be discovered in time by detecting the trust value. In this section, the capabilities of typical trust management schemes/models for defending against internal attacks are listed and analyzed as shown in Table 1.

The typical schemes for detecting and defending against the internal attacks with trust are summarized as follows:

**Denial of service attack:** after analysis, the power-aware trust model can effectively defend against DoS attacks (such as TMCE and DFDS); however, other trust management approaches based on event reporting will be affected by DoS attacks.

**Bad-mouthing attack/slander attack:** when defending against this attack, evaluation nodes can dynamically adjust the weight synthesized by indirect reputation according to the trust degree of neighbor nodes to mitigate the harm of slander attacks. Therefore, if the trust scheme only transmits positive information from other nodes, it can effectively defend against such attacks. In addition, the trust approach based on direct neighbor node trust perception or the scheme of multiple behavior observation aggregations is better able to defend against the slander attack. Moreover, GTMS, ReTrust, TDDG, LDTS, BTRES, BTMS, CRIT, HTMP, and so on can also defend against this attack.

**Ballot stuffing attack:** the confederate node of the malicious node improves reputation node by providing a large amount of successful interaction information to the other party. It is necessary to reduce the weight of the indirect trust value provided by the neighbor node in order to deal with such attacks. RFSN and ReTrust can defend against such attacks because of indirect trust values account for a small proportion in them.

**Collusion attack:** the attack requires more than one malicious node to cooperate, in order to provide the normal node wrong recommended value. Collusion attacks are more destructive, such as RFSN and GTMS can defend against the attack. In general, the trust model based on the direct observation of each node is not easily affected by collusion attacks. However, all of the other approaches of trust calculation are seriously jeopardized by collusion attacks. In defending against collusion attacks, nodes can set a threshold to filter out the indirect evaluation that is too different from direct evaluation to defend against collusion attacks.
Table 1: Defending against internal attack capability with trust management technology.

(a) Internal attacks | Trust management schemes
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADCT</td>
<td>ANFIS-TMM</td>
<td>TSN</td>
<td>BDTMS</td>
<td>BTMS</td>
<td>BTRES</td>
<td>CRIT</td>
<td>DFDI</td>
<td>DFR</td>
<td>EDTM</td>
<td>ETMRM</td>
<td>FANG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad-mouthing attack (slander attack)</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ballot stuffing attack</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black hole attack</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collusion attack</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflicting behavior attack</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data forgery attack</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denial of service attack</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garnished attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node replication attack</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-off attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reputation time-varying attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective forwarding attack</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selfish attack</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sinkhole attack</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleeper attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sybil attack</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: For ease of reference, the names of internal attacks and typical trust management techniques (abbreviations) in this table are arranged in ascending order of their initials. "√" indicates that the trust management scheme can defend against such internal attacks. "+" indicates that the trust management scheme can mitigate the harm of internal attacks or can only detect such internal attacks. "-" indicates that the defense ability of the trust management scheme against the internal attacks is unknown. "×" indicates that the trust management scheme cannot defend against the internal attacks.

(b) Internal attacks | Trust management schemes
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GMM-BRSN</td>
<td>GTMS</td>
<td>HTCW</td>
<td>HTMP</td>
<td>HTMS</td>
<td>ISTMM</td>
<td>JANANI</td>
<td>KARTHIK</td>
<td>KULDEEP</td>
<td>LIUCHENG</td>
<td>LDTS</td>
<td>LWTM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad-mouthing attack (slander attack)</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ballot stuffing attack</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black hole attack</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collusion attack</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflicting behavior attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data forgery attack</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denial of service attack</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garnished attack</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node replication attack</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-off attack</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reputation time-varying attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective forwarding attack</td>
<td>√</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selfish attack</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sinkhole attack</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleeper attack</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal attacks</td>
<td>GMM-BRSN</td>
<td>GTMS</td>
<td>HTCW</td>
<td>HTMP</td>
<td>HTMS</td>
<td>ISTMM</td>
<td>JANANI</td>
<td>KARTHIK</td>
<td>KULDEEP</td>
<td>LIU-CHENG</td>
<td>LDTS</td>
<td>LWTM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sybil attack</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(c)

<table>
<thead>
<tr>
<th>Internal attacks</th>
<th>MdTMM</th>
<th>Multi-agent</th>
<th>NBBTE</th>
<th>PLUS</th>
<th>RARRM</th>
<th>RASHMI</th>
<th>ReTrust</th>
<th>RFSN(BRSN)</th>
<th>SNTUA</th>
<th>TDDG</th>
<th>THI-CHAI</th>
<th>TMBBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad-mouthing attack (slander attack)</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ballot stuffing attack</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Black hole attack</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Collusion attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Conflicting behavior attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>Data forgery attack</td>
<td>✓</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>Denial of service attack</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Garnished attack</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Node replication attack</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>On-off attack</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reputation time-varying attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Selective forwarding attack</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>Selfish attack</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>×</td>
<td>✓</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Sinkhole attack</td>
<td>×</td>
<td>✓</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sleeper attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sybil attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>X</td>
</tr>
</tbody>
</table>

(d)

<table>
<thead>
<tr>
<th>Internal attacks</th>
<th>TMCED</th>
<th>TMM</th>
<th>TP-BRSN</th>
<th>TRTMS</th>
<th>Trust management schemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad-mouthing attack (slander attack)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>Ballot stuffing attack</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>×</td>
</tr>
<tr>
<td>Black hole attack</td>
<td>×</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>×</td>
</tr>
<tr>
<td>Collusion attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
<td>×</td>
</tr>
<tr>
<td>Conflicting behavior attack</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Data forgery attack</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Denial of service attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Garnished attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Node replication attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>On-off attack</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Reputation time-varying attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Selective forwarding attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Selfish attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>-</td>
</tr>
<tr>
<td>Sinkhole attack</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Table 1: Continued.
**Sleeper attack:** malicious nodes that act accurately in a certain period create a good reputation for themselves, and then be misbehaving. The aging scheme was introduced effectively to defend against such attacks in RFSN.

**On-off attack:** in on-off attack, malicious nodes perform sometimes well and sometimes poorly. Malicious nodes can maintain trust values even when they perform poorly. In order to cope with switching attacks, behavioral observations long ago cannot have the same aging weight as recent behavioral observations. Therefore, it can effectively defend against the on-off attack by using the trust approach of the forgetting factor. In this approach, the aging weight of behavioral observations long ago is lighter than the recent behavioral observations. In addition, it can also only use the current behavior observation to calculate the trust of the sensing node to defend against switching attacks. Therefore, TRTMS, FANG, PLUS, ReTrust, CRIT, and so on can effectively defend against the on-off attack.

**Selfish attack:** the self-node will simply delete the request and will not reserve the resource to send the trust reply after receiving the trust request. TDDG and others can effectively ensure network security through management technology increasing trust value.

**Garnished attack and reputation time-varying attack:** the behavior of a malicious node may be good or bad; the purpose is to remain undiscovered and cause damage. For example, when they accumulate a high degree of reputation, malicious nodes may attack suddenly. For garnished attacks, LDTS can defend against it, and for reputation time-varying attacks, TRTMS can effectively defend against it.

**Sybil attack:** ID authentication and centralized trust model are the good approaches to defend against Sybil attack, which can effectively identify the node and can also detect multiple false identities of the malicious node through the network sink node/BS.

**Conflict behavior attack:** considering that malicious nodes display different characteristics for different nodes, like defending the slander attack, conflict behavior attacks can be defended by trust approaches based on direct neighbor sensing (such as ATSN and TTSN) or aggregate multifactor observations (such as DFDI, TMBBT, and CTRT).

**Information attacks such as selective forwarding attack and data forgery attacks:** it is possible to obtain error information through the trust model just based on communication behavior, which makes the evaluation of reputation untrustworthy, and trust models or trust management schemes that effectively monitor all data forwarding and data integrity can defend against those attacks well.

**Sinkhole attack:** the attacker sets up a false aggregation node so that all information in the area “flows” to the false sink node. HTMP can defend against this attack.

**Node replication attack:** since the security credentials of the replicated nodes are cloned from the captured nodes, these replicated nodes can all be considered legitimate members of the network. Similar to the malicious nodes that launching Sybil attack, this type of replication attack by malicious nodes can also manipulate recommendations and elevate themselves as trusted nodes. Therefore, node replication attacks can be defended by ID verification (such as DFDI) and centralized trust model (such as GTMS), and BS can detect false identity.

The trust management schemes are mainly used to defend against the internal attacks, and different schemes aim at different internal attacks based on the requirements of applications. In addition, the trust value can be taken as a tool to solve the security issues for routing protocol in WSNs, due to the lower computational overhead. For hierarchical WSN, the cluster head is generally considered security, and it acts as a CA to provide the secure third-party recommendation. This can achieve the real-time of trust management. In planar WSNs, the trust decision is made by the cooperation between few neighbor nodes. The latter is suitable for those applications that are not real-time.

## 5. Future Directions in Trust

With the development of WSN, more and more researchers are paying attention to the trust management and proposing many novel trust models, schemes, and algorithms for WSN. However, the state-of-the-art studies in the field are still in the preliminary stage. In this section, we envision few potential research opportunities in the field as follows.

### 5.1. Trust Management System Based on Energy Efficiency

The limitation of various resources, including energy supplement, computational capabilities, and storage spaces, is a critical feature of WSN. Among which, the restriction on energy is one of the most important factors that restrict large-scale and long-term deployment of WSN. However, the existing trust management systems tend to require larger computation and additional communication energy consumption emerging from the interaction of some trust parameters, which will inevitably affect the lifetime of the network. On the other hand, effective analytical scheme for energy efficiency is still a blank in existing trust management systems. Therefore, it is of great importance to further investigate the energy-efficient trust management system and establish analytical scheme of energy-efficient system that
owns a certain objective evaluation basis. Meanwhile, in the designing process of trust management system, considering fully energy consumption and optimizing trust evaluation scheme are needed in order to improve the performance of the system.

5.2. Risk Evaluation Scheme for Trust Management. It is suggested that the risk evaluation scheme should be introduced and combined with trust management to establish a risk evaluation scheme oriented to trust management. The WSN is highly application-oriented, and the demand for trust management differs according to the different applications. Risk and trust factors should be taken into account when making decisions in different application environments. For example, the risk of nodes is compromised is different in military and home; the threshold of trust value can be adjusted properly under different risk levels to make the trust management system more stable, practical, and flexible.

5.3. Multiobjective Joint Optimization Mechanism for Node Information Forwarding with Trust. Considering the node’s trust value as a constraint condition, it is suggested that introduced it into the node information forwarding mechanism. For WSN, how to select the secure next hop is concerned with security, transmission, and energy consumption. Hence, through the analysis and evaluation of the multiobjective joint optimization method, a trade-off between the trust value, energy level, and transmission performance of neighbor nodes can be designed into a secure forwarding scheme for resource-constrained node. This is to defend against the internal attacks effectively and avoid deploying those high-strength encryption algorithms simultaneously.

Furthermore, the numerical size of trust value gradually becomes linguistic variable from a single decimal and then presents multidimensional matrix form. Apart from some specific dimensions, such as energy, which can be set artificially, other dimensions generally hazily input the unquantifiable dimensions through linguistic variables, such as outputting different grades of trust level through D-S theory. For data fusion in different dimensions, multidimension data was quantified by using multiple-input (including matrix form) single-output algorithms, such as the Analytic Hierarchy Process (AHP) and Gray Decision Model.

6. Conclusions
Although there have been many studies on trust, there is no concentrated research for WSNs. In this article, we systematically survey the research progress of the trust management processes and existing trust management in WSNs. Although trust management technology of traditional network is relatively mature, it cannot be directly applied to the resource-limited systems, such as WSN. Some existing trust management schemes in WSNs improve node security at the expense of other performances of the network, which may lead to the sacrifice of WSN lifetime.

Trust management in WSNs needs to meet the requirements as follows: WSN is a real-time network, so it must have low latency. The cost of memory, computing, and energy are expected to be minimized due to the limitations of sensor nodes’ own conditions. Through the analysis of existing trust management technology, further research and optimization are needed into the trust management scheme/system of WSN with the help of traditional network trust management scheme combing with specific application scenarios, especially the changes of wireless channels, the impact on trust valuation, and decision-making are fully considered. In view of this, we will gradually introduce energy efficiency, risk evaluation, and note mobility, as constraints in future work, and carry out research on efficient management scheme based on trust management combined with energy efficiency.

Conflicts of Interest
The authors declare no conflict of interest.

Acknowledgments
This work was supported in part by the National Natural Science Foundation of China (No. 51874300 and No. 61471346), the Shanxi Provincial People’s Government Jointly Funded Project of China for Coal Base and Low Carbon (No. U1510115), the Shanghai Natural Science Foundation (No. 17ZR1429100), the Science and Technology Innovation Program of Shanghai (No. 17511105903), the Fundamental Research Funds for State Key Laboratory of Synthetical Automation for Process Industries (No. PALN201703), the Scientific Instrument Developing Project of the Chinese Academy of Sciences (No. YJKYYQ20170074), the Open Fund Project of Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University (No. MJUKF-IPIC201905), and the National Key Research and Development Program of China—Internet of Things and Smart City Key Program (No. 2019YFB2101600, No. 2019YFB2101602, and No. 2019YFB2101602-03).

References


