jet INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2019, VOL. 65, NO. 2, PP. 259–266
Manuscript received November 24, 2018; revised April, 2019. DOI: 10.24425/ijet.2019.126309

# Secure Data Aggregation Mechanism based on Constrained Supervision for Wireless Sensor Network

Yubo Wang, Liang Li, Chen Ao, Puning Zhang, Zheng Wang, and Xinyang Zhao

***Abstract*—The data aggregation process of wireless sensor networks faces serious security problems. In order to defend the internal attacks launched by captured nodes and ensure the reliability of data aggregation, a secure data aggregation mechanism based on constrained supervision is proposed for wireless sensor network, which uses the advanced LEACH clustering method to select cluster heads. Then the cluster heads supervise the behaviors of cluster members and evaluate the trust values of nodes according to the communication behavior, data quality and residual energy. Then the node with the highest trust value is selected as the supervisor node to audit the cluster head and reject nodes with low trust values. Results show that the proposed mechanism can effectively identify the unreliable nodes, guarantee the system security and prolong the network lifetime.**

***Keywords*—Wireless Sensor Network, Security data aggregation, Trust management, Constrained supervision**

## I. INTRODUCTION

WIRELESS sensor network (WSN) is a multi-hop ad hoc network [1] composed of a large number of intelligent sensors distributed in the observation area. The sensor nodes in WSN cooperate with each other to perceive and collect the related data from monitored objects in the observation area, then upload the data step by step after simple processing to the sink node or base station. Because of flexible networking and simple deployment, WSN has been widely used in many fields such as national defense, medical treatment, environmental monitoring and so on [2],[3]. However, the sensor nodes are strictly resource restrained and deployed in a relatively harsh environment. The extremely limited storage, computing power and node energy inevitably restrict the further development of the WSN [4].

The rapid development of data aggregation [5] has solved the development bottleneck of WSNs. The sensor nodes are always densely deployed in the observation area, the data collected by different nodes during the same observation period often has high redundancy or strong correlation. The transmission of redundant data will increase the calculation and storage loads, and heavily waste network energy. Data aggregation can integrate the sensing data in the observed area, remove the redundant data and reduce the data complexity. Furthermore, data aggregation algorithms can greatly alleviate the data processing and storage loads of nodes, improve the data collection and transmission efficiency, and reduce the energy consumption of nodes and system [6]. However, sensor nodes are usually deployed in an open environment, which is vulnerable to malicious attacks and incurs security problems during the data aggregation process [7],[8],[9].

In view of the above problems, a Node Constrained Supervision (NCS) based secure data aggregation mechanism is proposed in this paper for WSNs. The main contributions of this paper are as follows.

1) We consider the residual node energy and use the advanced LEACH protocol to select cluster heads. The cluster head supervises the cluster members. According to the communication behaviors, data quality and residual energy, it selects the node with the highest trust value as the supervisor node to audit the cluster head and reject nodes with low trust values.

2) We guarantee the security of supervisor node and improve the network robustness. Internal attacks initiated by malicious nodes can be constrained by the supervision among common nodes, cluster heads and supervisor nodes. .

3) When supervising node behaviors, we consider the residual energy to balance the energy consumption of nodes and effectively avoid the premature death of nodes with high reliability.

The rest of this paper is organized as follows. Some related works are discussed in Section II. The network structure is presented in Section III. We introduce the clustering method in Section IV. We describe the trust evaluation method in Section V. The data aggregation algorithm is proposed in Section VI. The experiment results are given in Section VII. Finally, we conclude this paper and present our future work.

## II. RELATED WORK

By capturing sensor nodes or even colluding cluster heads, malicious attackers launch internal attack in WSN to intercept the aggregated data. Once the aggregated data is tampered, the final sensor result obtained by the sink will suffer a grave impact, which misleads the user to make wrong judgment. In addition, some malicious nodes also send forged data to some sensor nodes frequently, which causes the network congestions, resource wastes, increased transmission delay

This work was supported by the grant No. 546816180001 financed from State Grid Corporation Headquarters Science and Technology Project.

Y.B. Wang, L. Li, C. Ao and Z. Wang are with State Grid Key Laboratory of Power Industrial Chip Design and Analysis Technology, Beijing Smart-Chip Microelectronics Technology Co., Ltd., Beijing, China.

P.N. Zhang is with Chongqing University of Posts and Telecommunications, Chongqing, China.(e-mail: zhangpn@cqupt.edu.cn).

X.Y. Zhao is with Maintenance Company of State Grid NingXia Electric Power Co.,Ltd, Yinchuan,China.

and premature energy depletion of sensor nodes. Therefore, an efficient data aggregation mechanism should be designed to defend malicious attacks, which ensures the reliability of aggregated data and prolongs the network lifetime.

Aiming at the security of data aggregation in WSN, researchers proposed various solutions. The traditional mechanisms encrypted the aggregated data against tampering attacks. Paper [10] proposed a secure data aggregation mechanism based on the Hash Tree, which was a symmetric encryption mechanism. In addition, the Concealed Data Aggregation (CDA) mechanism was proposed in paper [11] to divide data into several blocks randomly and data blocks were aggregated after being multiplied by a key, which was a homomorphic encryption mechanism. The encryption based data aggregation mechanism was easy to implement but usually had high computational complexity, which was not suitable for WSNs with limited computing power. Besides, the keys were vulnerable due to the compromising of nodes. Therefore, with the escalation of attack means, the encryption based data aggregation mechanism fails to guarantee the data security. To this end, some researchers proposed supervision based data aggregation mechanisms, [12],[13] which can select supervisor nodes to audit the node behaviors and identify malicious nodes. In addition, paper [14] and [15] introduced trust management mechanisms into data aggregation. In the process, the trust value of node was evaluated according to the communication behaviors, then the supervision mechanism was designed to secure the data aggregation. However, existing supervision based data aggregation mechanisms monitor node behaviors without considering the residual node energy, which results in excessive use of some nodes, causes the premature energy depletion of these nodes and severely affects the network lifetime. Besides, these mechanisms do not consider the security of supervisor nodes. Once a supervisor node is captured or colludes with attackers, the network security will be seriously threatened.

## III. NETWORK STRUCTURE

Fig. 1 is the architecture diagram of WSN. The entire observation area is divided into several clusters. In the proposed mechanism, the nodes can be divided into four categories according to their functions: cluster member, cluster head, supervisor node and sink node.

Cluster member: they are randomly deployed in the observation area, perceive and collect the related data of the monitored objects, and then send the sensor data to the cluster head for data aggregation. There are multiple cluster members in each cluster.

Cluster head: each cluster only has one cluster head, and the data collected by the cluster members is aggregated according to a certain data aggregation algorithm. Aggregated data is sent to the sink node through relay nodes.

Supervisor node: the node with the highest trust value is selected as the supervisor node to audit the behaviors of the cluster head. Each cluster has one supervisor node and the supervisor node does not participate in the data collection process.

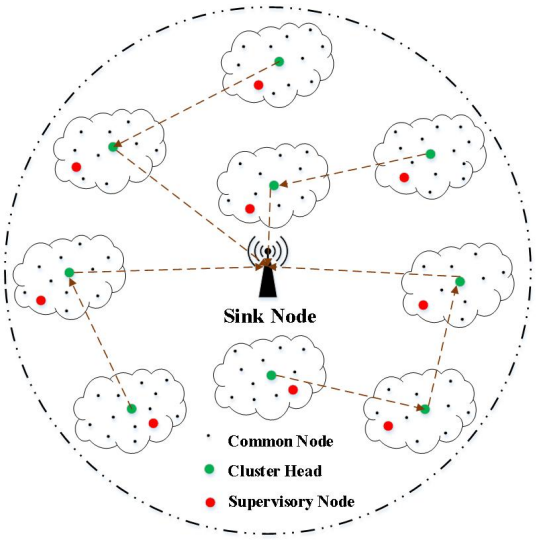


Fig. 1. Architecture diagram of wireless sensor networks.

Sink node: the WSN has one sink node, which processes data uploaded by cluster heads and sends the processed data to the user via Internet or satellite. The sink node has more storage and stronger computing ability. In general, the base station acts as the sink node, which can be completely trusted.

Fig. 2 is the sketch map of the proposed mechanism, which periodically selects cluster head and supervisor node. Each runtime cycle is called one round and contains four processes: (1) Network clustering. At the end of each cycle, all nodes become cluster members, then the cluster head is selected according to the advanced LEACH mechanism at the beginning of the next cycle. The cluster head broadcasts a message to cluster members in proximity. Cluster members select their own clusters to join according to the signal strength of cluster heads. (2) Supervisor node selection. According to the trust value obtained from the last round, the cluster member with the highest trust value in the cluster is selected as the supervisor node to audit the behaviors of the cluster head. (3) Sensing task execution. After determining the cluster head and supervisor node, the cluster members in the cluster begin to perform sensing tasks, collect and send the sensor data to the cluster head. The cluster head aggregates the data after evaluating the received data and uploads the aggregated data to the sink node. (4) Trust value update. The cluster head updates the trust values of cluster members according to their communication behaviors, data quality and residual energy. The trust value of the cluster head is evaluated by the supervisor node. The quality of the data uploaded by cluster head is assessed by the sink node, the sink node feeds back evaluation results to the supervisor nodes. The dynamic update of trust values and the periodic reselection of cluster heads and supervisor nodes can effectively detect malicious nodes, guarantee the reliability of aggregated data and enhance the system security. Besides, constrained supervision on nodes can enhance the security of cluster heads and supervisor nodes, while preventing malicious attacks.

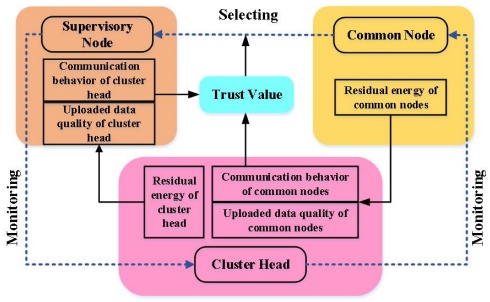


Fig. 2. Sketch map of the proposed mechanism.

## IV. NETWORK CLUSTERING

We use the advanced Low-Energy Adaptive Clustering Hierarchy(LEACH) mechanism to select cluster heads and divide the network into clusters. The basic idea of the traditional LEACH mechanism [16] is to determine the optimal number of clusters K according to the number of nodes in the network. Then each node in the network generates a random number between [0, 1]. If the random number is less than the preset threshold $T(n)$, the node is selected as the cluster head. The threshold $T(n)$ is calculated based on the optimal cluster number K and the total number of nodes N. The LEACH mechanism periodically changes cluster heads to balance the energy consumption of nodes and prevent the premature energy depletion of cluster heads. The threshold $T(n)$ calculated by the traditional LEACH clustering mechanism ensures that the cluster head in the last period will not be selected as the cluster head of the current cycle, so that each node is not continuously selected as a cluster head, which can reduce the damage to the network due to the capture of cluster heads. But only involving the number of former cluster heads to evaluate the residual energy of nodes is inaccurate and cannot balance the energy consumption of nodes in the network.

In order to balance the energy consumption of nodes and prolong the network lifetime effectively, we improve the threshold evaluation method of the traditional LEACH mechanism. Residual energy parameter E and energy consumption parameter D are employed in the improved threshold evaluation. The residual energy parameter $E = E_c/E_{in}$, where $E_c$ is the residual energy of a given node and $E_{in}$ is the initial energy of nodes. Most energy of nodes is consumed for data forwarding and the cluster head is responsible for forwarding the aggregated data to the sink node. Therefore, we use the shortest hop number $H_{\min}$ between a given node and the sink node to evaluate energy consumption parameter D, that is $D = H_{\min}$. Finally, threshold $T(n)$ in the advanced LEACH mechanism is obtained:

$$T(n) = \begin{cases} \sqrt{E^2(n)/D(n)} \times \frac{P(n)}{1-P(n)\times[rmod(1/P(n))]} n \in G \\ 0 \text{otherwise} \end{cases} \quad (1)$$

where $P(n) = K/N$ denotes the probability of a node being selected as a cluster head, $K$ is the optimal number of clusters obtained according to paper [16], $N$ is the total number of nodes in the WSN, $r$ is the current network cycle and $G$ is the set of nodes that are not selected as cluster heads in the past $1/P(n)$ cycles. The cluster heads broadcast the selection results in the network, then other nodes choose their own clusters according to the signal strength of the received message to complete the network clustering.

Using the advanced LEACH mechanism to change cluster heads periodically can balance the energy consumption of nodes and prolong the network lifetime. Meanwhile, it also prevents the same node from being continuously selected as the cluster head, which effectively reduces the damage of the captured cluster heads.

## V. TRUST EVALUATION

The trust value is quantified for data security and node reliability based on the communication behaviors and attributes of nodes[17]. In this paper, the trust management is integrated into the data aggregation. We prejudge whether a node will initiate attacks by evaluating its trust value. We update trust values after each round of sensing task and reject nodes with low trust values, so as to ensure the security and stability of the network. Meanwhile, the node with the highest trust value is selected as the supervisor node to realize the constrained supervision among all nodes, which secures the cluster head and supervisor node and prevents them from malicious attacks.

### A. Trust attribute quantification

The node trust value is comprehensively evaluated from three aspects: interaction trust, data trust and energy trust.

The interaction trust reflects the communication behaviors of nodes. In WSN, malicious node behaviors mainly include data interception, tampering and retransmission. The CSMA/CA protocol of the data link layer describes the node communication process as follows. When node A sends data to node B, node B needs to feedback a confirmation message to node A upon receiving the data and node A monitors the confirmation message from node B. If A receives the confirmation message, the data communication between A and B is successful. Otherwise, it is failed. Excluding channel factors, the communication failure is mainly caused by the malicious behaviors of nodes in most cases [18]. Therefore, monitoring the communication behaviors of nodes to quantify the interaction trust can detect malicious nodes within the network.

When node i initiates communication with node j, the result only contains two cases: communication success or failure. The numbers of successful communications and failed communications can be used to calculate the interaction trust $Ctr_{i,j}$ between node i and j. $Ctr_{i,j}$ increases with the rising number of successful communications and decreases with the growing number of communication failures. $Ctr_{i,j}$ is normalized into . According to its trend, we quantify the interaction trust between nodes using the Beta distribution of the number of successful and failed communications:

$$Ctr_{i,j}(T_n) = E\left(Beta(Sn_{i,j}(T_n)+1, fn_{i,j}(T_n)+1)\right) \times \frac{1}{\sqrt{fn_{i,j}(T_n)}} = \frac{Sn_{i,j}(T_n)+1}{Sn_{i,j}(T_n)+fn_{i,j}(T_n)+2} \times \frac{1}{\sqrt{fn_{i,j}(T_n)}} \tag{2}$$

Where $E(*)$ denotes the average of and being multiplied by $1\Big/\sqrt{fn_{i,j}(T_n)}$ can attenuate the interaction trust sharply when the number of failed communications increases. Finally, the interaction trust of nodes i can be evaluated by the average interaction trust between node i and other nodes within the whole cluster:

$$Ctr_i(T_n) = E(Ctr_{i,j}(T_n)) \tag{3}$$

Data trust reflects the quality of uploaded data. Malicious nodes may tamper or forge the aggregated data and send the same or similar data frequently when launching attacks. These data generally has bad quality and low trust value. We obtain the data trust value of nodes by verifying the consistency of uploaded data.

Data consistency mainly refers to the difference between uploaded data of nodes from the same cluster. In order to accurately distinguish reliable data uploaded by trustworthy nodes and false data uploaded by malicious nodes, then further to evaluate the trust values of nodes, we use the Cloud theory to verify the consistency of the uploaded data. The Cloud theory [19] combines fuzziness and randomness perfectly to map the qualitative and quantitative relationship. As an uncertain transformation model of qualitative and quantitative combination, the Cloud theory can solve the consistency verification of data as an information processing problem.

In the monitored area, the data of a certain sensing task collected by node n during one cycle is recorded as $d_n$, then the sensor data set collected by the cluster head is denoted by:

$$D(T_n) = \{d_1, d_2, ..., d_n\} \tag{4}$$

We calculate the expectation $E_{D(T_n)}$ and variance $S_{D(T_n)}$ of all the sensor data, then the entropy of the sensor data set $En_{D(T_n)}$ is calculated by $E_{D(T_n)}$:

$$En_{D(T_n)} = \sqrt{\frac{\pi}{2}} \times \frac{1}{n} \sum_{i=1}^{n} \left| d_i - E_{D(T_n)} \right| \tag{5}$$

Next, we calculate the hyper entropy of sensor data set $He_{D(T_n)}$ according to $En_{D(T_n)}$ and $S_{D(T_n)}$:

$$He_{D(T_n)} = \sqrt{S^2_{D(T_n)} - En^2_{D(T_n)}} \tag{6}$$

Finally, we obtain a random normal function with expectation $E_{D(T_n)}$ and variance $S_{D(T_n)}$, and then we use the cloud membership degree of sensor data to denote the data trust of node i:

$$Dtr_i = \sigma_{D(T_n)}(d_i) = \exp\left(-\frac{\left(d_i - En_{D(T_n)}\right)}{2He^2_{D(T_n)}}\right) \tag{7}$$

The cloud membership degree reflects the attribution degree of the data collected by a node to the overall sensor data set. When the cloud membership degree of a node is very low, the data uploaded by this node is deviant and very possible to be false or falsified, which means this node is abnormal and its data trust value is low.

At last, we quantify the energy trust value of nodes and reject nodes with insufficient energy to avoid the premature death of these nodes and network congestion. According to the residual energy of nodes, the energy trust value of nodes can be evaluated by:

$$Etr_i = E_i^c / E_i^{in} \tag{8}$$

where $E_i^c$ is the current residual energy of node and $E_i^{in}$ is the initial energy of node.

### *B. Comprehensive trust estimation*

At the end of each cycle, the cluster head combines the interaction, data and energy trust to update the trust value of nodes:

$$Tr_i = \omega_1 \times Ctr_i + \omega_2 \times Dtr_i + \omega_3 \times Etr_i \tag{9}$$

Where $\omega_1, \omega_2, \omega_3$ denote the weights of interaction, data and energy trust respectively. The traditional mechanisms usually set a fixed trust weight according to actual scenarios. However, in practical applications, the network environment of WSN is constantly changing. The fixed trust weight cannot provide the dynamic adaptability. As the network runs and environment changes, it is difficult to ensure the evaluation accuracy of trust values. Entropy is the parameter describing the uncertainty of objectives. The information entropy reflects the amount of information contained in the trust value. Therefore, we use the information entropy theory [20] to evaluate the amount of information contained in the evaluation results of trust, so as to dynamically and reasonably determine the trust weights.

We calculate the information entropy of the interaction, data and energy trust of nodes within a cluster, as shown by Eq. (10-12):

$$H_{Ctr} = -\sum_{i=1}^{n} Ctr_i \cdot \log_2(Ctr_i) \tag{10}$$

$$H_{Dtr} = -\sum_{i=1}^{n} Dtr_i \cdot \log_2(Dtr_i) \tag{11}$$

$$H_{Etr} = -\sum_{i=1}^{n} Etr_i \cdot \log_2(Etr_i) \tag{12}$$

Then we allocate the trust weights according to the amount of information contained in the trust values, as shown by Eq. (13):

$$\omega_k = \begin{cases} \frac{H_{Ctr_i}}{H_{Ctr_i}+H_{Dtr_i}+H_{Etr_i}} k{=}\,1 \\ \frac{H_{Dtr_i}}{H_{Ctr_i}+H_{Dtr_i}+H_{Etr_i}} k{=}\,2 \\ \frac{H_{Etr_i}}{H_{Ctr_i}+H_{Dtr_i}+H_{Etr_i}} k{=}\,3 \end{cases} \tag{13}$$

At last we combine Eq. (13) and Eq. (9) to obtain comprehensive trust value $Tr_i$ of node i.

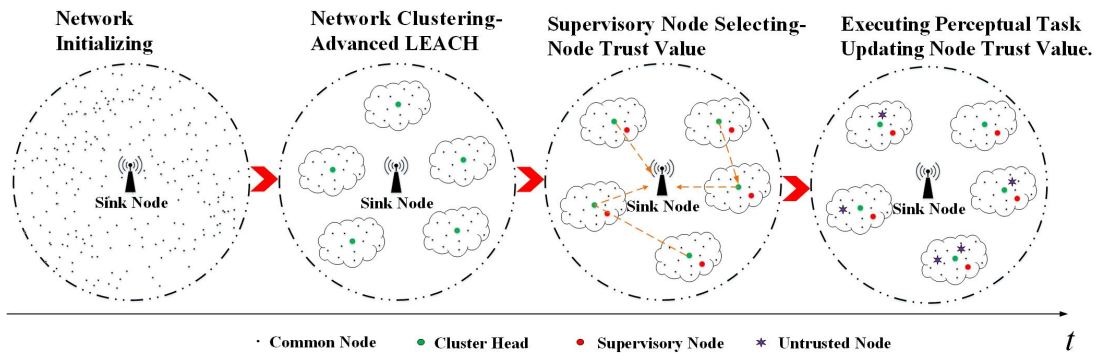


Fig. 3. Flowchart of the proposed mechanism.

## VI. ALGORITHM DESCRIPTION

Fig. 3 is the flowchart of the proposed mechanism. The specific steps are as follows:

Step 1. Network initializing: trust values of all nodes are set to 1 and all nodes are completely trusted.

Step 2. Network clustering: we select cluster heads and divide the entire observation area into clusters by adopting the advanced LEACH clustering mechanism.

Step 3. Supervisor nodes selecting: we choose the node with the highest trust value within a cluster as the supervisor node in this cycle to monitor the cluster head.

Step 4. Sensing task executing: the cluster members collect data and upload them to the cluster head. After the cluster head collects the sensor data, it verifies the data consistency to update the data trust values of nodes within the cluster, then aggregates the uploaded data and sending the aggregated data to the sink node. Sink node verifies the data uploaded by the cluster head and feedbacks the verification result to the supervisor node to update the data trust value of the cluster head.

Step 5. Trust value updating: at the end of each cycle of sensing tasks, the cluster head updates the trust value of the cluster members and the supervisor node updates the trust value of the cluster head. Finally, the supervisor node rejects the untrustworthy nodes in the cluster according to the updated trust values.

The trust value of malicious or untrustworthy nodes is usually far lower than the normal node, which can be regarded as an abnormal value. With the help of abnormal value detection, we can reject the untrustworthy nodes with low trust values. In this paper, we use the Grubbs criterion [21] to verify the trust value of nodes in the cluster at the end of each cycle to solve the problem of multiple untrustworthy nodes coexisting in a cluster. Then we can identify and reject the untrustworthy nodes to ensure the reliability of aggregated data and the performance and stability of WSNs.

During the $n$th cycle, the cluster head evaluates trust values of cluster members and the supervisor node evaluates the trust value of the cluster head after performing the sensing task. Then we sort the trust values of all nodes in descending order to form the set of trust values $Tr\left(T_n\right)$. Second, we calculate expectation $E_{Tr(T_n)}$ and variance $S_{Tr(T_n)}$ of trust values of all nodes. As described earlier, the abnormal trust value is usually small in the set of trust values, so the node with the minimum trust value is considered as the suspect node firstly. Then we employ the difference between the suspect trust value and $E_{Tr(T_n)}$ to get the suspect trust value offset. Next, we calculate $G_i$ of the suspect trust value according to Eq. (14):

$$G_i = \frac{E_{Tr(T_n)} - \min\left(Tr\left(T_n\right)\right)}{S_{Tr(T_n)}} \tag{14}$$

At last we compare $G_i$ of the suspect trust value with critical value $G_p\left(n\right)$ given by the Grubbs table. If $G_i$ of the suspect trust value is larger than the critical value, this node is untrustworthy. Then we continue to select the minimum trust value after rejecting the untrustworthy node and repeat the above steps until $G_i$ of the suspect trust value is smaller than the critical value.

## VII. NUMERICAL RESULTS

In order to verify the feasibility and effectiveness of the proposed secure data aggregation mechanism, we use the NS2 simulation platform to evaluate the security performance and network lifetime of the proposed NCS mechanism. The contrast algorithms are the Trust Management based security data aggregation Mechanism(TMM) proposed in [14] and the Risk Analysis based security data aggregation Mechanism(RAM) proposed in [15]. During the simulation process, 120 sensor nodes are deployed in the area of $120m \times 120m$ and the communication range of nodes is 10m. The initial energy of all nodes is set to 3J, the packet size is 256bits, the number of cluster heads is 6 and the energy consumption of sending a bit data is 50nJ/bit. The maximum number of network cycles is 100 and the initial trust value of nodes is set to 1. The attack behaviors of malicious nodes include tampering data, replay attack and Denial of Service attack.

### A. *Comparative analysis of network security*

In this part, validity of the proposed trust evaluation method is simulated and compared with the other two algorithms. The simulation results are shown in Fig. 4 and 5.

Fig. 4 shows the influence of the proportion of untrustworthy nodes in the network on the discovery probability of

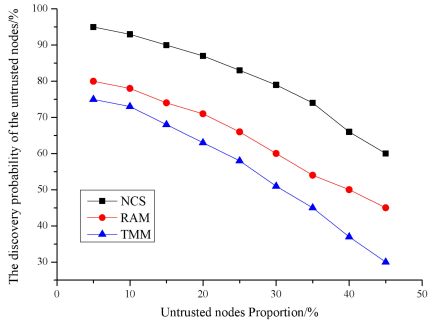


Fig. 4. Influence of the proportion of untrustworthy nodes on the discovery probability of untrustworthy nodes.

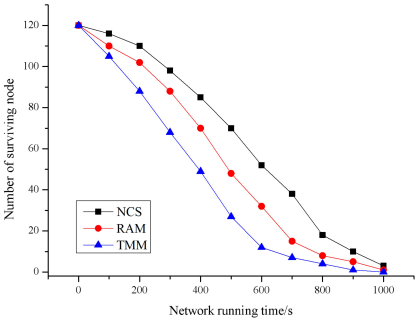


Fig. 6. Variation of the number of surviving node with the network runtime.

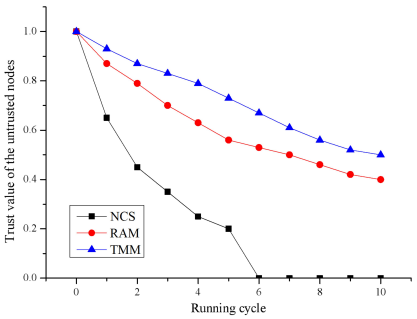


Fig. 5. Trend of the trust value of untrustworthy nodes.

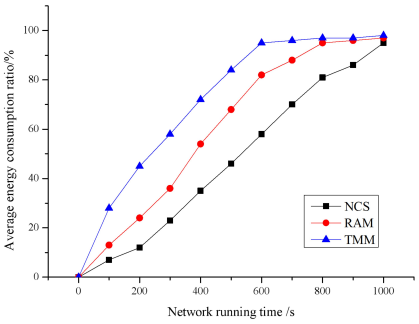


Fig. 7. Variation of the average energy consumption ratio.

untrustworthy nodes. It can be seen that the proposed mechanism has a 20% higher discovery probability of untrustworthy nodes, because that the proposed mechanism considers the interaction, data and energy trust to evaluate the node trust comprehensively and uses dynamic weights to adjust the trust value according to the actual evaluation results. It can accurately reflect the behavioral characteristics of the untrustworthy nodes, so that the untrustworthy nodes can be detected easier. The TMM only considers the communication behaviors of nodes, its trust value estimation accuracy is the worst and the discovery probability of untrustworthy nodes is the lowest. The RAM considers both communication behaviors and data quality of nodes, it has the improved performance than TMM. However, RAM does not consider the residual energy of nodes, neglects the untrustworthy nodes of insufficient energy, and its detection performance is still inferior to our proposed mechanism. The discovery probability of untrustworthy nodes of the three mechanisms decreases with the growing proportion of untrusted nodes, because untrustworthy nodes affect the normal communication of the WSN, consume too much system resource and have a severely negative impact on the algorithm performance.

Fig. 5 shows the trust value of untrustworthy nodes in the network changes with the network runtime. It can be seen that the trust value of untrustworthy nodes in the three mechanisms decreases with the network runtime. The proposed trust evaluation method has better sensitivity than the other two mechanisms, because our mechanism considers the impact of abnormal behaviors on the trust value evaluation result. At the same time, the designed node supervision mechanism can evaluate and detect abnormal node behaviors timely. Once malicious or abnormal behaviors occur, the trust value of this node will drop sharply. High sensitivity of the proposed trust evaluation method can minimize the impact of abnormal node behaviors on the network performance and reject untrustworthy nodes timely. After the fifth network cycle, the trust value of untrustworthy nodes is reduced to 0 by the proposed mechanism, because the node with low trust value has been rejected by the system.

Fig. 4 and Fig. 5 show that the proposed mechanism collects the behavior and state information of nodes accurately through the constrained supervision among nodes. By evaluating the trust value of nodes comprehensively and detecting the untrustworthy nodes in time, it can ensure the security and stability of the system and the accuracy of the aggregated data.

### *B. Comparative analysis of network lifetime*

In this part, validity of the proposed trust evaluation method is simulated and compared with the other two algorithms. The simulation results are shown in Fig. 4 and 5.

Fig. 6 shows the variation of numbers of surviving nodes in the network as the network runs. In three mechanisms,

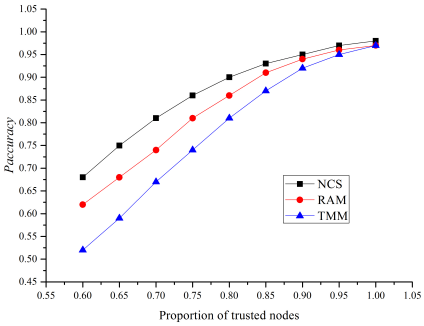


Fig. 8. Influence of the proportion of trusted nodes on the data aggregation accuracy.

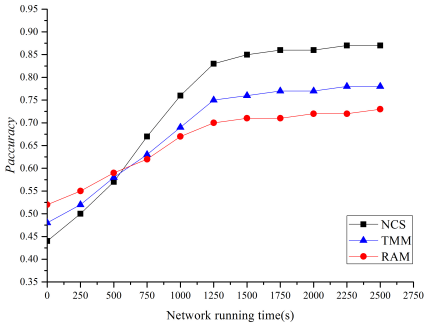


Fig. 9. Variation of the data aggregation accuracy with the network runtime.

the node energy is gradually exhausted and the number of surviving node gradually reduces as the network runs. The node death rate of our mechanism is low and shows the obvious advantage over the other two mechanisms, because that the proposed mechanism considers the residual energy of nodes in both the cluster head selection process and trust value evaluation process, which effectively avoids the premature death of the overused reliable nodes. At the same time, malicious nodes are rejected in time, which can prevent malicious nodes from launching replay attacks to consume the node energy. The proposed algorithm prolongs the survival time of normal nodes and improves the network lifetime greatly.

Fig. 7 shows the variation of the average energy consumption ratio of the network with the network runtime. The average energy consumption ratio of the proposed mechanism is always lower than those of the other two mechanisms, because the other two mechanisms do not consider the residual energy of nodes, resulting in the excessive use and the premature energy depletion of many reliable nodes. After a node dies, its neighbor nodes will become isolated, then they look for other paths to transmit data, which consumes more energy. On the other hand, the other two mechanisms cannot detect malicious nodes effectively and malicious nodes in the network will also waste more energy which results in the excessive energy consumption and shortens network lifetime.

### C. Comparative analysis of data aggregation accuracy

We compare and analyze the data aggregation accuracy of the three algorithms, and the simulation results are shown in Fig. 8 and Fig. 9. Data aggregation accuracy is calculated by Eq. (15):

$$P_{accury} = 1 - \frac{|f - r|}{r} \tag{15}$$

where $f$ denotes the data aggregation result and $r$ is the actual value.

Fig. 8 presents the influence of proportions of trusted nodes on the data aggregation accuracy. The data aggregation accuracy of the three mechanisms increases with the growing proportion of trusted nodes. However, under the same proportions of trusted nodes, the proposed mechanism can evaluate the trust value of nodes according to multiple factors, detect the untrustworthy nodes accurately, and provide higher data aggregation accuracy. Fig. 9 shows the variation of data aggregation accuracy with the network runtime. It can be seen that the data aggregation accuracy of the three mechanisms gradually increases and tends to be stable with the network runtime. The TMM and RAM do not consider the residual energy of nodes, which causes the overload of reliable nodes, so that their data aggregation accuracy is slightly higher. As the network runs, the reliable nodes die too early because of the energy depletion, which isolates many nodes and lowers their data aggregation accuracy.

## VIII. Conclusion

Aiming at the hidden security risks in the data aggregation and the shorten network lifetime of WSNs, a constrained supervision based secure data aggregation mechanism is proposed. we consider the residual energy of nodes and employ the advanced LEACH to cluster the sensor nodes. Then a trust management mechanism considering the node communication behaviors, data quality and residual energy is designed to evaluate the node trust comprehensively. According to the trust value evaluation, the supervisor nodes are selected to form the constrained supervision system among nodes, which can detect the abnormal malicious behaviors of nodes and reject untrustworthy nodes in time. Results show that the proposed mechanism can secure the data aggregation process, prolongs the network lifetime and provides accurate and reliable aggregated data. In future, we will further analyze the node behavior characteristics and consider more factors when evaluating the trust value of nodes in order to prevent new types of attacks.

## References

[1] D. P. Wu, J. He, H. G Wang, C. G. Wang and R. Y. Wang, "A hierarchical packet forwarding mechanism for energy harvesting wireless sensor networks", *IEEE Communication Magazine*, 53 (8), 92–98 (2015).

[2] H. Barani, Y. Jaradat, H. Huang, Z. C. Li and S. Misra, "Effect of sink location and redundancy on multi-sink wireless sensor networks: a capacity and delay analysis", *IET Communications*, 12 (8), 941–947 (2018).

[3] P. N. Zhang and J. Ma, "Channel Characteristic Aware Privacy Protection Mechanism in WBAN", *Sensors*, 18 (8), accepted, (2018). DOI: 10.3233/JCS-2007-15104

[4] D. Upadhyay, A. Dubey and P. Thilagam, "Application of non-linear gaussian regression-based adaptive clock synchronization technique for wireless sensor network in agriculture", *IET Communications*, 18 (10), 4328–4335 (2018).

[5] D. P. Wu, B. R. Yang, H. G. Wang, C. Y. Wang and R. Y. Wang, "Privacy-preserving multimedia big data aggregation in large-scale wireless sensor networks", *ACM Trans. Multimedia Computing*, 12 (4), 1–19 (2016).

[6] S. Boubiche, D. Boubiche, A. Bilami and H. Toral-Cruz, "Big data challenges and data aggregation strategies in wireless sensor networks", *IEEE Access*, 6 (C), 20558–20571 (2018).

[7] Al-Tabbakh, S. M, "Novel technique for data aggregation in wireless sensor networks", 2017 International Conference on Internet of Things, October 2017, Gafsa, Tunisia, 2017, pp. 1–8.

[8] P. N. Zhang, X. Y. Kang, Y. Z. Liu, and H. P. Yang, "Cooperative Willingness aware Collaborative Caching Mechanism towards Cellular D2D Communication", *IEEE ACCESS*, accepted, (2018). DOI: 10.1109/ACCESS.2018.2873662

[9] P. N. Zhang, X. Y. Kang, D. P. Wu, and R. Y. Wang, "High-accuracy entity state prediction method based on deep belief network towards IoT search", *IEEE Wireless Communications Letters*, accepted, (2018). DOI: 10.1109/LWC.2018.2877639

[10] H. Chan, A. Perrig, B. Przydatek and D. Song, "SIA: Secure information aggregation in sensor networks", *Journal of Computer Security*, 15 (1), 1–19 (2007).

[11] D. Westhoff, J. Girao and M. Acharya, "Concealed data aggregation for reverse multicast traffic in sensor networks: encryption, key distribution, and routing adaptation", *IEEE Transactions on Mobile Computing*, 5 (10), 1417–1431 (2006).

[12] X. Dong and S. Li, "A secure data aggregation approach based on monitoring in wireless sensor networks", *China Communication*, 9 (6), 14–17 (2012).

[13] H. Wang, Y. Li, M. R. Mi and P. Wang, "Secure data fusion method based on supervisory mechanism for Industrial Internet of Things", *Chinese Jounal of Scientific Instrument*, 34 (4), 817–824 (2013).

[14] K. Yong, Y. Xin, "Study on trust management-based cluster-head selection in wireless sensor networks", 2015 4th International Conference on Advanced Information Technology and Sensor Application (AITS), August 2015, Harbin, China, 2015, pp. 1–8.

[15] N. Labraoui, "A reliable trust management scheme in wireless sensor networks", 2015 12th International Symposium on Programming and Systems (ISPS), April 2015, Algiers, Algeria, 2015, pp. 1–6.

[16] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks", 33rd Annual Hawaii International Conference on System Sciences, January 2000, Maui, USA, 2000, pp. 223.

[17] D. P. Wu, H. P. Zhang, H. G. Wang, C. G. Wang, R. Y. Wang and Y. Xie, "Quality of Protection (QoP)-driven data forwarding for intermittently connected wireless networks", *IEEE Wireless Communication*, 22 (4), 66–73 (2015).

[18] D. P. Wu, S. S. Si, S. E. Wu, and R. Y. Wang, "Dynamic trust relationships aware data privacy protection in Mobile Crowd-Sensing", *IEEE Internet of Things Journal*, 5 (4), 2958–2970 (2018).

[19] D. Li, Y. Du, "Uncertain artificial intelligence", Beijiing: National Defense Industry Press, 2005, pp. 137–185.

[20] L. Qin, K. Q. Sun and S. G. LI, "Maximum fuzzy entropy image segmentation based on artificial fish school algorithm", International Conference on Intelligent Human-Machine Systems and Cybernetics, December 2016, Hangzhou, China, 2016, pp. 164–168.

[21] W. Luo, Y. Wu, J. Yuan, W. Lu, "The calculation method with Grubbs test for real-time saturation flow tate at signalized intersection", The Second International Conference on Intelligent Transportation, Singapore, 2017, pp. 129–136.