Research Article

A Data Aggregation Privacy Protection Algorithm Based on Fat Tree in Wireless Sensor Networks

Cheng Li ^{1,2}, Guoyin Zhang ¹, Yan Mao,¹ and Xin Zhao¹

¹College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

²College of Computer and Control Engineering, Qiqihar University, Qiqihar 161006, China

Correspondence should be addressed to Guoyin Zhang; zhangguoyin@hrbeu.edu.cn

Received 12 March 2021; Accepted 11 August 2021; Published 13 September 2021

Academic Editor: Mohamed Amine Ferrag

Copyright © 2021 Cheng Li 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.

Wireless sensor network is the momentous part of the Internet of Things. Data aggregation technology is the most practical method to reduce the amount of communication among nodes. Adding a privacy protection mechanism to data aggregation is one of the important means for privacy protection and security in wireless sensor networks. Aiming at certain performance defects of the existing SMART (Slice-Mix-AggRegaTe) privacy protection algorithms, a fat tree-based data aggregation privacy protection algorithm is proposed in this paper, which is referred to as the FTSMART (Fat Tree Slice-Mix-AggRegaTe) algorithm. Concerning the innovative algorithm, the fat tree (FT) is introduced, the fat tree structure is adopted to optimize the data slicing scheme and the aggregation tree generation scheme, and the allocation of fixed time intervals is employed for nodes to reduce the transmission collision in the data aggregation process and to guarantee the completion of the data transmission. The simulation experiment results demonstrate that the FTSMART algorithm has presented the favorable performance in terms of the privacy protection, the network communication overhead, and the data aggregation accuracy.

1. Introduction

Wireless sensor network (WSN) is the multihop and self-organized networks formed from a large amount of wireless sensor nodes by wireless communications. A large number of sensor nodes are randomly deployed in various environments (such as harsh environments where humans cannot stay for a long time) to perceive and collect the target information so that people can analyze and process the information and make the reasonable judgments. Currently, wireless sensor network has been widely applied in the environmental monitoring, the military fields, the intelligent transportation, the logistics tracking, the intelligent medical, and so forth. According to the energy-constrained characteristics of wireless sensor network, the data aggregation technology is widely used. The core of the data aggregation technology is to aggregate the data from different information sources, remove the redundant information, and reduce the amount of transmitted data by means of the data compression and the feature extraction. Accordingly, the

network energy consumption is decreased, the network life cycle is extended, and the efficiency and accuracy of data collection are improved greatly [1].

Wireless sensor network has a low security to some certain. In the meantime, the data privacy protection mechanism is not provided by the basic data aggregation technology generally. However, in practical applications, due to the characteristics of the wireless transmission, the data transmitted among nodes is easy to be captured and eavesdropped, and the data of the child nodes is obtained by the trusted parent nodes in networks. If the wireless link is broken or the parent nodes are captured by the attacker, the private data is exposed accordingly. The data aggregation privacy protection technology of wireless sensor network is the technology that prevents the private data from being acquired even if the transmitted data is captured and decrypted externally or captured by other trusted nodes internally, under the condition that the data aggregation result is correct. So far, in the previous researches, some privacy protection schemes have been proposed, each of

which has its own scope of application, and some of the schemes still have some problems that need to be further resolved.

Based on the research of the data slicing aggregation in the SMART (Slice-Mix-AggRegaTe) algorithm [2], the fat tree [3] is introduced into the domain of wireless sensor network, and a fat tree-based data aggregation privacy protection algorithm is proposed in this paper. An aggregation tree is constructed through the fat tree, a hop-by-hop data aggregation method is adopted, and the data slicing technology is used for the privacy protection. The main contributions of the paper are as follows:

- (1) The fat tree structure is introduced to construct a fusion tree. The structure and characteristics of the fat tree are fully utilized, and the constructed fusion tree is the shortest path tree, which realizes the advantages of low computational complexity, less data transmission, and small fusion delay.
- (2) The fat tree structure is combined with the data slicing technology to protect data privacy. The data is sliced according to the number of parent nodes in the fat tree where the fusion tree node is located, and the sensing data of all the nodes is sliced and fused, which improves the performance of data privacy protection.
- (3) The fixed time intervals (time slices) are allocated to complete the data fusion scheduling. The process of generating the fusion tree from the fat tree is carried out layer by layer from the leaf node to the root node. At the same time, a fixed time interval can be allocated to each node according to the scheduling principles, so as to avoid the data transmission collision among nodes and improve the accuracy of fusion.

The remainder of the paper is organized as follows. Section 2 reviews the previous works. In Section 3, we described the FTSMART algorithm, including the system model and the specific implementation process in detail. In Section 4, the performance of the FTSMART algorithm is evaluated by the simulation experiments, covering the privacy protection and the communication overhead as well as the aggregation accuracy. Section 5 concludes the research and the future work.

2. Previous Works

He et al. [2] proposed the Privacy-preserving Data Aggregation (PDA) algorithm and researched the additive aggregation function SUM, which included the Cluster-based Private Data Aggregation (CPDA) algorithm and the SMART algorithm. But the CPDA algorithm was the data aggregation algorithm based on clustering, and its calculation process was complicated and computationally expensive. The SMART algorithm is an algorithm closely related to this paper. Its central thoughts are to cut the acquired data into several fragments (i.e., slices), and to transmit the sliced data along different transmission paths. In this way, unless

all the slices are obtained by the privacy attacker, the ultimate private information cannot be obtained. The specific implementation process of the SMART algorithm is divided into 3 steps:

In the first phase (slicing), the collected data is randomly sliced into J slices by each node, 1 slice is kept by itself, and the remaining $(J-1)$ data slices are encrypted and randomly sent to the neighboring nodes.

In the second phase (mixing), the intermediate node waits for a period of time to receive the data slices sent by other nodes. When the intermediate node receives the encrypted data slice, the shared key is used to perform decryption, and then all the data slices perform the mixed calculation.

In the final phase (aggregating), all the nodes employ the data aggregation tree that are established by the Tiny Aggregation service for ad hoc sensor networks (TAG) algorithm [4] to transfer up the data aggregation results layer by layer.

The advantages of the SMART algorithm are that each node slices its own private information and mixes the data slices of different nodes, which increases the difficulty for an attacker to eavesdrop the complete data multiplication. It can be indicated that the algorithm has the prominent performance of the data privacy protection as well as the less computational overhead. The disadvantages of the SMART algorithm are that, with the high network communication overhead, the number of data packets generated is several times that of many other algorithms, and it is not suitable for the networks with a large number of nodes and the densely distributed networks. Otherwise, the amount of the sliced data is larger, which not only affects the normal working efficiency of the networks but also seriously affects the accuracy of the data aggregation.

A great amount of research studies have been carried out on the SMART algorithm by many scholars, and plenty of improvements have been presented [5–13]. The study in [5] proposed the EEHA algorithm, which only permitted the data collected by the leaf nodes of the aggregation tree to be sliced, so that the number of slices was reduced, and the network communication overhead was reduced, but the data privacy protection performance was reduced as well. The study in [6] presented the ESMART algorithm on the basis of [5]. The number of the leaf node data slices was random between $[2, J]$ (J is the maximum number of slices that were set), which improved the data privacy protection performance greatly. In [7], the HEEPP algorithm was proposed on the basis of [6], which added a data query mechanism, queried the data transmission status of the subnodes when idle, prevented the data loss, and thus improved the aggregation accuracy. In [8], in order to improve the accuracy of the aggregation, the LTPART algorithm was presented. In the data aggregation phase, a fixed time slice or a floating time slice was allocated to each layer of nodes to guarantee the full aggregation. In [6] and [7], there existed the problem that the number of the random slices was larger than the number of the neighboring nodes, so the slicing data could not be exchanged and mixed normally. In order to solve the problem, the study in [9] introduced the SESDA algorithm,

and the number of packets was dynamically determined by the number of the data slices that the nodes had received. The ESPART algorithm proposed in [10] reduced the data packets generated by segmentation and reassembly by controlling the inside and outside of nodes. The study in [11] put forward the PSMART algorithm and added a local slicing strategy. The nodes that failed to be sliced have not undergone slicing, but end-to-end encryption. In order to further improve the performance of the data privacy protection, the work in [12] proposed the D-SMART algorithm, and the data was divided according to the deviation degree, the ordinary data was cut into 2 slices, the important data was cut into 3 slices, and the confidential data was cut into 4 slices. Aiming at the data collision problem caused by the segmentation and recombination technology, the work in [13] introduced five optimization factors to reduce the collision rate and reduce the loss caused by collision. Some scholars have proposed various related applications based on the SMART algorithm, which have been illustrated in [14, 15] as well.

On the basis of the related researches above, an aggregation tree that is suitable for the fragmented data aggregation is constructed to research the privacy protection algorithm in the paper, and the SMART algorithm and its relative algorithms have been improved to a certain extent. Meanwhile, the energy consumption overhead performance of network communications is taken into account.

3. Data Aggregation Privacy Protection Algorithm Based on Fat Tree

The fat tree-based data aggregation privacy protection algorithm is proposed in the paper, which is composed of four steps: the fat tree construction phase, the slicing phase, the mixing phase, and the aggregating phase. In this section, the system model, the implementation process, and the advantages of the proposed algorithm are elaborated and illustrated in detail.

3.1. System Model. In a two-dimensional square area of $L \times L$, there are N nodes randomly distributed, and only one base station is considered, namely, the Sink node, which constitutes the extremely connected wireless sensor network $G(V, E)$ with Sink. All the nodes in the network G receive and send data and have the same transmission radius r . In the research on the problem, the following conditions are set:

- (1) All the nodes in the network have completed the data perception.
- (2) The data aggregation that occurs in the network is the complete aggregation; that is, no matter how many information is received by node i , the node encapsulates all the information in a data packet for transmission. The data aggregation function is defined [16, 17], which is shown in

$$y(t) = f(d_1(t) + d_2(t) + \dots + d_i(t)), \quad (i = 1, 2, \dots, N), \quad (1)$$

where $d_i(t)$ represents the data collected by node i at time t . The research in the paper supports the additive aggregation calculation, such as summation, expectation, or variance. A typical SUM aggregation function is shown in

$$y(t) = \sum_{i=1}^N d_i(t). \quad (2)$$

According to formula (2), the aggregation process using the SUM function is shown in Figure 1.

- (3) It is agreed upon that any node in the network sends and receives the data within 1 hop, but sending data and receiving data are nonparallel, and receiving two or more data packets at the same time is nonparallel as well.
- (4) All the transmissions that can be carried out at the same time (no collision between each other) are completed in a unit of time, which is recorded as a time interval. The internode transmissions have the uniform time intervals.
- (5) As with the SMART algorithm, a random key distribution mechanism is adopted to encrypt and decrypt the transmitted data in the paper and complete the communications among nodes through a shared key [2, 12]. The probability that two random nodes in the network have the same key is expressed as follows:

$$Peavesdrop = \frac{k}{K}, \quad (3)$$

where K is the total number of keys in the key pool and k is the number of keys in the key ring. *Peavesdrop* represents the probability that the links between any pair of communication nodes can be cracked by the attacker, that is, the probability that the private data is eavesdropped on. When the key pool is large enough, the security of the mechanism is stronger.

3.2. Fat Tree Construction Phase. The purpose of constructing the fat tree is to serve as the basis of node data slicing, finally cutting it into the shortest path aggregation tree. Therefore, the characteristics of the fat tree are listed as follows: there is a unique root node, namely, the Sink node, to ensure that the data is finally gathered on the Sink node; the distance between any node and its parent nodes (except the Sink node) or between any node and its child nodes (except for the leaf nodes) must be less than or equal to the node communication radius r , that is, the one-hop distance, to ensure that data is not lost; the path from any node to the root node is not unique, but the number of hops through these paths is the same; that is, the depth of the node is unique to ensure that the path from any node to the root node is the shortest path, which is conducive to reduce the time delay of the aggregation process.

FIGURE 1: Schematic diagram of aggregation of SUM function.

At this phase, the specific process of constructing the fat tree on the randomly generated node set (the point set of wireless sensor network G) is as follows: firstly, the center point of G is selected as the Sink node S , which is the root node of the fat tree; then, the Sink node is used as the base node, and the node communication radius r is used as a hop distance to search for its child nodes, and the child nodes obtained are linked to the fat tree; then, the child nodes obtained are applied as the base node, and the node communication radius r is applied as the one-hop distance to search for its child nodes (the nodes that are not in the fat tree), and the process is looped until all the nodes are connected to the fat tree. In this way, the fat tree is successfully constructed, and the process of searching for the child nodes is shown in Figure 2.

In Figure 2, the child nodes searched by the Sink node S employing communication radius r are nodes 1, 2, and 3; the child nodes searched by node 1 employing communication radius r are nodes 4 and 5; the child nodes searched by node 2 employing communication radius r are nodes 5 and 6; the child nodes searched by node 4 employing the communication radius r are nodes 7 and 8, and so forth. Then, child nodes 1, 2, and 3 are linked to node S , nodes 4 and 5 to node 1, and nodes 5 and 6 to node 2, until all the nodes are linked, and the fat tree is constructed successfully. The fat tree structure is presented in Figure 3.

3.3. Slicing Phase. All the nodes in the network G cut their own sensory data into $n + 1$ slices according to the number n of their parent nodes in the fat tree (the Sink node is not considered); retain 1 slice of them randomly, encrypt the rest of the data slices, and send to its parent nodes randomly. The schematic diagram is shown in Figure 4.

In Figure 4, node 5 is taken as an example to illustrate the data slices. Node 5 has three parent nodes 1, 2, and 3 in the fat tree. Therefore, the sensory data of node 5 is cut into 4 slices, namely, r_{51} , r_{52} , r_{53} , and r_{54} .

In that case, all the nodes in the network G have performed the sensory data slicing, which greatly improves the performance of the data privacy protection; in addition, the number of the parent nodes in the fat tree is known, which prevents the blind slicing and reduces the failure rate of slicing the data communication.

FIGURE 2: The process of searching for child nodes.

FIGURE 3: Fat tree structure.

3.4. Mixing Phase. After all the nodes in the network G receive the encrypted data slices, the shared key is used to perform decryption, and then a mixing calculation is performed on these data slices as well as the retained data slices. The mixed calculation process is illustrated in Figure 5. The SUM function is applied to carry out the data slice mixing calculation, as shown in

$$A_j = \sum_{i,j \in U_j} r_{ij}. \quad (4)$$

A_j is the new data packet after the mixing calculation in node j . U_j is the collection of node j and its child nodes i in the fat tree, as well as the collection of node j and node i that sends the data slice to node j . r_{ij} means the data packet of the data slice from node i to node j .

In Figure 5, node 5 is taken as an example to describe data mixing. Node 5 has three data slices before mixing, namely, slice r_{55} that is reserved by itself, data slice r_{85} received from node 8, and data slice r_{95} received from node 9.

FIGURE 4: Schematic diagram of data slices.

FIGURE 5: Schematic diagram of data mixing.

Node 5 performs the mixing calculation according to the SUM function to obtain the new data A_5 after mixing.

3.5. Aggregating Phase. A shortest path tree in the fat tree is selected as the aggregation tree, and after the mixing calculation, the new data packet is transferred up from all the leaf nodes to the Sink node layer by layer. In this process, after each node receives all the decrypted data, it performs a data aggregation calculation, and then the data encryption continues to be transmitted to the upper layer. According to the selected aggregation tree structure, the fixed time interval (i.e., time slice, the time to complete one-hop data transmission) is allocated to the node to ensure the completion of the data transmission. The aggregation scheduling process is shown in Figure 6.

The network starts the data transmission at time t_0 . After t time intervals, all the information is transmitted to the Sink

node, and the aggregation cycle ends. Then, the aggregation cycle of the network is t (the data transmission time is much longer than the aggregation calculation time, and the set time interval is slightly longer than the data transmission time). As in Figure 6, the aggregation cycle of network G is 4.

Therefore, the aggregation cycle of the network G is fixed, which ensures the completion of data aggregation transmission, prevents the data package loss, and improves the aggregation accuracy.

4. Experimental Results and Analysis

The simulator embedded in TinyOS is used as a simulation tool to conduct simulation experiments on the privacy protection performance of the SMART algorithm [2], the D-SMART algorithm [12], and the proposed FTSMART algorithm. The simulation experiments incorporate the comparison experiment of data privacy protection performance, the comparison experiment of communication overhead, and the comparison experiment of aggregation accuracy.

4.1. Simulation Environment. In the simulation environment, the wireless sensor network is deployed as follows: 500 sensor nodes are randomly distributed in a two-dimensional rectangular area of 500 m * 500 m. The specific parameter settings are shown in Table 1.

4.2. Privacy Protection Performance. After the nodes in the network use the slicing technology, if an attacker attempts to attain the data of a certain node, all the access links of this node must be cracked to restore its original data. $P(q)$ is defined as the probability that the private data of a node is decrypted, and $P(q)$ is used as a measure of the privacy protection, where q represents the probability that the link among the nodes is decrypted.

For the SMART algorithm, the probability of cracking the node privacy data is shown in

$$P(q) = q^{J-1} \sum_{k=0}^{d_{in_max}} P(\text{in-degree} = k) q^k, \quad (5)$$

where J expresses the number of the node data slice, q^{J-1} expresses the probability that the out-degree link of the node is cracked, k expresses the number of the in-degree links, d_{in_max} expresses the maximum number of the in-degree links of the node and is determined by the number of the neighboring nodes, $P(\text{in-degree} = k)$ expresses the probability that the in-degree link is equal to k , and q^k expresses the probability that the in-degree link of the node is cracked. Therefore, the SMART algorithm is obviously affected by the number J of the node data slices.

For the D-SMART algorithm, the probability of cracking the node privacy data is shown in

FIGURE 6: Schematic diagram of data aggregation scheduling.

TABLE 1: Simulation parameter settings.

Parameter	Implication	Value
N	Number of nodes	500
L	Simulation area (m^2)	$500 * 500$
r	Communication radius (m)	25
Pa	Proportion of fusion nodes	20%

$$P(q) = \sum_{j \in \{1,2,3,4\}} P(\text{out-degree} = j) q^{j-1} * \sum_{k=0}^{d_{in_max}} P(\text{in-degree} = k) q^k, \quad (6)$$

where j represents the number of the out-degree links. For the leaf nodes, $j \in \{2, 3, 4\}$. For other nodes, $j = 1$, $P(\text{out-degree} = j)$ represents the probability that the in-degree link is equal to j , q^{j-1} indicates the probability that the out-degree link of the node is cracked. Therefore, the D-SMART

algorithm is obviously affected by the aggregation node ratio Pa .

For the FTSMART algorithm, the probability of cracking the node privacy data is shown in the following formula:

$$P(q) = \sum_{j=1}^{d_{out_max}} P(\text{out-degree} = j) q^{j-1} * \sum_{k=0}^{d_{in_max}} P(\text{in-degree} = k) q^k, \quad (7)$$

where d_{out_max} signifies the maximum number of the out-degree links of the node and is determined by the number of the parent nodes in the fat tree, the direct child nodes j of the Sink node is equal to 1, d_{in_max} is determined by the number of the direct child nodes in the fat tree, and only the leaf node k is equal to 0. Therefore, the FTSMART algorithm is affected by the fat tree structure to some extent.

The simulation experiment results of the privacy protection performance of the SMART algorithm, the D-SMART algorithm, and the proposed FTSMART algorithm are demonstrated in Figure 7.

It can be observed from Figure 7 that as the probability of cracking the communication link increases, the probability of the privacy data exposure of the SMART algorithm, the D-SMART algorithm, and the proposed FTSMART algorithm also increases monotonically. Among them, the

FTSMART algorithm has the lowest probability of the privacy data exposure. The D-SMART algorithm only has the leaf nodes for the dynamic slice. As the proportion of the aggregation nodes Pa decreases, the privacy protection capability of the D-SMART algorithm exceeds that of the SMART algorithm. The FTSMART algorithm, like the SMART algorithm, slices all the node data. The SMART algorithm slices each node data and satisfies $J = 3$, and as J increases, the privacy protection capability is enhanced. The number of the data slices for each node of the FTSMART algorithm is determined by $n + 1$ slices according to the number n of their parent nodes in the fat tree. The distribution and density of the nodes in this experimental environments determine that the FTSMART algorithm has the highest privacy protection capability than the SMART algorithm and the D-SMART algorithm.

FIGURE 7: Comparison of data privacy protection performance.

4.3. Data Communication Overhead. In the wireless sensor network, the node energy consumption mainly derives from the data transmission among nodes. A large amount of the data transmission result in the premature death of some nodes, thus shortening the entire life cycle of the network. We adopt the total amount of the data packets transmitted in the network as a measure of the communication overhead. The simulation experiments are carried out to compare the communication overhead of the SMART algorithm, the D-SMART algorithm, and the FTSMART algorithm.

For the SMART algorithm, in the slicing phase, each sensor node cuts the sensory data into J slices and sends $(J-1)$ slices to the neighboring nodes to generate $(J-1)$ data packets. In the aggregating phase, each sensor node sends a new data packet after mixing to the upper node. Therefore, the communication overhead of the SMART algorithm is expressed in

$$O_c = N * (J + 1) + N * 1 = N * J, \quad (8)$$

where O_c signifies the communication overhead and N signifies the total number of the nodes in the network.

For the D-SMART algorithm, in the slicing phase, the aggregation node data does not need to be sliced. The leaf node data is dynamically sliced. In the aggregating phase, each sensor node needs to send one data packet to the upper node. Therefore, the communication overhead of the D-SMART algorithm is expressed in

$$O_c = N * Pa + \sum_{i=1}^{N * (1-Pa)} j_i (j_i \in \{2, 3, 4\}), \quad (9)$$

where $(1-Pa)$ represents the proportion of the leaf nodes and j_i represents the number of the data slices of node i , that is, the total number of the data packets generated by node i .

For the FTSMART algorithm, in the slicing phase, all the sensor nodes need to slice the data into $(n+1)$ slices according to the number n of their parent nodes. In the

aggregation phase, each sensor node needs to send one data packet to the upper node. Therefore, the communication overhead of the D-SMART algorithm is expressed in

$$O_c = \sum_{i=1}^N j_i (j_i \in [1, 2, \dots, n_{\max} + 1]), \quad (10)$$

where i represents the node sequence number and $n_{\max}$ represents the maximum number of the parent nodes for all the sensor nodes.

The simulation experiment results of the communication overhead of the SMART algorithm, the D-SMART algorithm, and the proposed FTSMART algorithm are demonstrated in Figure 8.

It has been observed from Figure 8 that the network communication overhead of the SMART algorithm in different aggregation cycles is the highest, maintaining 1500 data packets. The network communication overhead of the D-SMART algorithm varies from 1100 to 1300 data packets. And the network communication overhead of the FTSMART algorithm is the lowest, maintaining 1100 data packets. In the SMART algorithm, the number of the node slices is fixed ($J=3$), and the communication overhead is also fixed. In the D-SMART algorithm, the aggregation node ($Pa=0.2$) does not participate in the slicing behavior. The leaf nodes change in the interval $[2, 4]$ according to the number of different slices of the sensory data, and the communication overhead fluctuates as well. All the nodes of the FTSMART algorithm participate in the slicing behavior, and the number of the slices depends on the number of the parent nodes in the fat tree. The fat tree structure in the network is fixed, and the communication overhead is fixed accordingly.

4.4. Data Aggregation Accuracy. In wireless sensor network, the data aggregation accuracy is one of the important metrics that demonstrate the performance of the data aggregation algorithms. In theory, the accuracy of the aggregation result is 100%, which signifies that the ultimate aggregation result is equal to the sum of the data collected by each node in the network [6]. However, in practical applications, the data transmission collision, the delay, the bit error, and the packet loss in the data aggregation process are unavoidable, which mainly resulted from the use of shared channels to transmit data in wireless sensor network, thus making the accuracy of the aggregation result lower. The calculation of the data aggregation accuracy is illustrated in

$$P = \frac{D^*}{\sum_{i=1}^N D_i}, \quad (11)$$

where P denotes the data aggregation accuracy and D^* expresses the final aggregation result obtained by the Sink node. D_i represents the sensory data of node i , and $\sum_{i=1}^N D_i$ represents the sum of sensory data of all the nodes in the network.

The simulation experiment results of the aggregation accuracy of the SMART algorithm, the D-SMART algorithm, and the proposed FTSMART algorithm are demonstrated in Figure 9.

FIGURE 8: Comparison of communication overhead.

FIGURE 9: Comparison of aggregation accuracy.

It can be detected from Figure 9 that, within an aggregation cycle, as the time interval increases, the aggregation accuracy of these three comparison algorithms increases monotonically and the increase in speed becomes slower. The aggregation accuracy rates of the SMART algorithm, the D-SMART algorithm, and the proposed FTSMART algorithm are 70%, 83%, and 91%, respectively. Among them, the SMART algorithm has the lowest aggregation accuracy, and the FTSMART algorithm has the highest aggregation accuracy. Within the aggregation cycle, the more the data transmitted, the more the probability of collision in the transmission process, the more the data lost, and the greater the impact on the aggregation result. That is, an aggregation scheme with low communication overhead can obtain the favorable data aggregation accuracy. In the

FTSMART algorithm, the aggregation accuracy is improved by optimizing the aggregation tree and reasonably allocating the time intervals.

5. Conclusion

On the basis of analyzing and researching the SMART algorithm and the relative algorithms, the FTSMART algorithm is proposed, and the fat tree is introduced into the data aggregation of wireless sensor network, which has greatly improved the deficiencies of the SMART algorithm in the data privacy protection and the aggregation accuracy. From the simulation experiments, it has been detected that the FTSMART algorithm is affected by the fat tree structure to some extent, and therefore the process of optimizing the fat tree to generate the aggregation tree can be taken into consideration. The research of the paper has set the conditions of the equal time interval transmission and the complete aggregation, which requires further expansion. In the future, the researches on the multimodal data aggregation methods and the multimodal data security protection will be the huge challenges.

Data Availability

All of the data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] R. Ramesh and P. K. Varshney, "Data aggregation techniques in sensor networks: a survey," *Communications Surveys & Tutorials*, vol. 8, no. 4, pp. 48–63, 2006.
- [2] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher, "PDA: privacy-preserving data aggregation in wireless sensor networks," in *Proceeding of the 26th IEEE International Conference on Computer Communications*, pp. 2045–2053, IEEE Computer Society Press, Washington, DC, USA, 2007.
- [3] C. E. Leiserson, "Fat-trees: universal networks for hardware-efficient supercomputing," *IEEE Transactions on Computers*, vol. C-34, no. 10, pp. 892–901, 1985.
- [4] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: a Tiny Aggregation service for ad-hoc sensor networks," *Acm Sigops Operating Systems Review*, vol. 36, no. SI, pp. 131–146, 2002.
- [5] H. Li, K. Lin, and K. Li, "Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks," *Computer Communications*, vol. 34, no. 4, pp. 591–597, 2011.
- [6] C. Li and Y. Liu, "ESMART: energy-efficient slice-mix-aggregate for wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 9, no. 12, pp. 1–9, Article ID 134509, 2013.
- [7] C.-X. Liu, Y. Liu, Z.-J. Zhang, and Z.-Y. Cheng, "High energy-efficient and privacy-preserving secure data aggregation for wireless sensor networks," *International Journal of Communication Systems*, vol. 26, no. 3, pp. 380–394, 2013.

- [8] Q. Liang and Y. A. N. G. Geng, "A low traffic privacy-preserving aggregation algorithm," *Computer Technology and Development*, vol. 23, no. 8, pp. 133–136, 2013.
- [9] T. C. Wang, X. L. Qin, L. Liu et al., "Secure and energy-efficient spatial data aggregation algorithm in wireless sensor networks," *Journal of Software*, vol. 25, no. 8, pp. 1671–1684, 2014.
- [10] G. Yang, A.-Q. Wang, Z.-Y. Chen, J. Xu, and H.-Y. Wang, "An energy-saving privacy-preserving data aggregation algorithm," *Chinese Journal of Computers*, vol. 34, no. 5, pp. 792–800, 2011.
- [11] L. I. Sen and Y. A. N. G. Geng, "Research on precision aggregation privacy-preserving algorithm in wireless sensor networks," *Computer Technology and Development*, vol. 23, no. 9, pp. 139–142, 2013.
- [12] J. Wang and Y. Chen, "Research and improvement of wireless sensor network secure data aggregation protocol based on SMART," *International Journal of Wireless Information Networks*, vol. 25, no. 3, pp. 232–240, 2018.
- [13] G. Yang, S. Li, Z. Y. Chen et al., "High-accuracy and privacy-preserving oriented data aggregation algorithm in sensor networks," *Chinese Journal of Computers*, vol. 36, no. 1, pp. 189–200, 2013.
- [14] J. Shi, R. Zhang, Y. Liu et al., "PriSense: privacy-preserving data aggregation in people-centric urban sensing systems," in *2010 Proceedings IEEE INFOCOM*, pp. 758–766, San Diego, CA, USA, March 2010.
- [15] W. He, H. Nguyen, X. Liu et al., "iPDA: an integrity-protecting private data aggregation scheme for wireless sensor networks," in *Proceeding of MILCOM 2008 - 2008 IEEE Military Communications Conference*, pp. 1–7, San Diego, California, USA, November 2008.
- [16] V. Kumar and S. Madria, "PIP: privacy and integrity preserving data aggregation in wireless sensor networks," in *Proceeding of 2013 IEEE 32nd International Symposium on Reliable Distributed Systems*, pp. 10–19, IEEE, Braga, Portugal, September 2013.
- [17] K. Parmar and D. C. Jinwala, "Malleability resilient concealed data aggregation in wireless sensor networks," *Wireless Personal Communications*, vol. 82, no. 1, pp. 777–798, 2015.