Research Article

A Balance Privacy-Preserving Data Aggregation Model in Wireless Sensor Networks

Changlun Zhang, Chao Li, and Yi Zhao

Science School, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

Correspondence should be addressed to Changlun Zhang; zclun@bucea.edu.cn

Received 5 January 2015; Accepted 21 March 2015

Academic Editor: Qing-An Zeng

Copyright © 2015 Changlun Zhang 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 networks are always deployed in remote and hostile environments to gather sensitive information, in which sensor nodes are apt to encounter some serious leakage of sensitive data. Hence, privacy-preserving is becoming an increasingly important issue in security data aggregation for wireless sensor networks. In this paper, we propose a balance privacy-preserving data aggregation model (BPDA) based on slicing and mixing technology. Compared to fixed or random slicing, BPDA model gives a balance slicing mechanism to ensure that slice can be sent to the nodes which have lower privacy preservation and enhance the privacy-preserving efficacy. Furthermore, according to the influence of the node degree and energy, three different schemes are presented to keep the privacy-preserving data aggregation balance. Theoretical analysis and simulation show that BPDA model demonstrates a good performance in terms of privacy-preserving efficacy and communication overhead and prolongs the lifetime of network.

1. Introduction

A wireless sensor network (WSN) is a typical ad hoc network which is highly distributed and self-organized. It usually consists of plenty of small sensor nodes which gather the data from its monitoring physical or environment conditions (e.g., the temperature, the sound, etc.) and send their data to the destination (base station) directly or via multihop [1, 2]. WSN has many popular applications [3, 4], such as military surveillance, industrial process monitoring and control, air pollution monitoring, and machine health monitoring. Sensor node has typical weakness such as processing capability, storage capacity, and limited energy. In particular, the sensor nodes are always deployed in the harsh environment, without being recharged or replaced. Therefore, energy efficiency in in-network data processing is very important for WSN.

In WSN, sensor nodes collect regional information and upload them to the base station, where the base station disposes these data to obtain the result. There are plenty of redundant data in the process of uploading. For example, hundreds of sensor nodes are used to collect the temperature of an area while the manager just wants to know the maximum temperature. So, it is not necessary to send all the temperature data but a derivative such as maximum to base station. Data aggregation [5, 6] aims to aggregate redundant data at intermediate sensor nodes applying a suitable aggregation function on the received data. Aggregation reduces the amount of network traffic which helps to reduce energy consumption on sensor nodes.

WSN is always deployed in unsecured and untrusted environment, which makes it exposed to all kinds of intrusions, and encounters some serious security issue. Some works [7–12] studied security of data aggregate in WSN. These schemes use cryptographic mechanism to establish secure communication links for data aggregation. In some special scenario, the individual sensitive data should not be disclosed to any node in the network, including parent node or neighboring node. This is privacy-preserving [13, 14] in WSN, which keeps private data from being intercepted and used by adversaries and untrusted nodes and maintains data privacy of a sensor node from other trusted neighboring nodes in the WSN. Nowadays, privacy-preserving is becoming an increasingly important issue for security of WSN [15–24]. SMART (Slice-Mix-AggReTa) [18] is a typical scheme,
which slices individual sensitive data into a fixed set of pieces and sends them to corresponding associated nodes. Afterwards, some improved approaches, such as iPDA [19], PEPDA [23] and ESPART [22], were proposed.

In this paper, we propose a balance privacy preserving-data aggregation (BPDA) model for WSN. Our work focuses on the distribution mechanism of slice for privacy data, which considers balance slices distribution based on the random distribution. It reduces the redundancy of the privacy preservation efficacy and prolongs the lifetime of the WSN.

The remainder of this paper is organized as follows. In Section 2, the related work is summarized. In Section 3, preliminaries of our work are described. A balanced privacy-preserving data aggregation model is proposed in Section 4. Section 5 analyzes the privacy preservation efficiency of proposed schemes. Performance evaluation and analysis are described in Section 6. Finally, the conclusion of this paper is given.

2. Related Work

Recently, secure data aggregation is becoming an important issue for wireless sensor networks. Cryptograph has been an efficient mechanism to secure data aggregation. Generally, there are two typical encryption methods: end-to-end scheme and hot-by-hop scheme. End-to-end scheme [15–17] needs to establish secure link between base station and each sensor node before data transmission, and then encrypted data is transmitted to base station directly. Hot-by-hop scheme [18–23] needs sensor node to encrypt data before sending and decrypt them after receiving. The shortcoming of this scheme is that it cannot provide data confidentiality in the node during the process of decryption and encryption.

Some existed works on secure data aggregation focused on symmetric key cryptography to achieve end-to-end security. Recently, homomorphic encryption technique is introduced to achieve in-network aggregation, which allows the ciphertext to be aggregated directly, and then the receiver verifies if decrypted aggregation result matches the result of aggregation operations performed on plaintext. Castelluccia et al. [15] proposed a homomorphic encryption scheme based on addition operation named AHE. AHE is a simple and provably secure encryption scheme that allows efficient additive aggregation of encrypted data. Only one modular addition is necessary for ciphertext aggregation. CDA [16] is an approach that conceals sensitive data end-to-end but still provides efficient and flexible in-network data aggregation. The aggregating intermediate nodes are not able to read the sensitive plaintext data. Ozdemir and Xiao [17] proposed a novel integrity protecting hierarchical concealed data aggregation protocol, which employs an elliptic curve cryptography-based homomorphic encryption algorithm. The scheme can offer data integrity and confidentiality along with hierarchical data aggregation. In addition, during the decryption of aggregated data, the base station is able to classify the encrypted and aggregated data based on the encryption keys. But homomorphism based secure data aggregation schemes need more computation overhead, and they cannot be used in the network which is divided into plenty of clusters. These schemes were described to deal with addition operations in data aggregation with homomorphic encryption, such as finding sum or average value. Homomorphic encryption makes it possible to aggregate data without doing encryption and decryption at intermediate nodes. However, it is not easy to find out operation satisfying the homomorphic properties.

Meanwhile, a typical slicing technology is introduced into privacy-preserving data aggregation in WSN. He proposed SMART scheme [18] firstly which includes three steps of slicing, mixing, and aggregation. In slicing step, each node slices its private data into $J$ pieces randomly and keeps one of the $J$ pieces by itself while sending the remaining $J-1$ pieces to the neighbor nodes. Mixing step comes after all nodes finished slicing their own data. In mixing step, each node sums up all the slices which include the slices it has received and the one slice it kept. In the aggregation step, all nodes aggregate the data and send the result to the query server. The SMART scheme scatters the data over the neighbor nodes. The attackers must eavesdrop enough communication channels if he wants to obtain the data collected by some node. This makes the difficulty of eavesdropping increase radically. In [19], He et al. improved their scheme and presented iPDA scheme which is an integrity-protecting private data aggregation scheme. In iPDA, data privacy is still achieved through SMART scheme while data integrity is achieved through redundancy by constructing disjoint aggregation trees to collect data of interests. But it inherits the weakness of SMART—large communication overhead. Groat et al. [20] studied nonlinear aggregation functions instead of traditional additive function, and then presented $K$-indistinguishable privacy-preserving data aggregation (KIPDA) scheme which achieves the goal of privacy preserving upon MAX and MIN aggregation functions by obfuscating data being forwarded. Aiming at cutting down the large communication overhead, Liu et al. [21] improved the process of the slicing and proposed a high energy-efficient and privacy-preserving (HEEPP) secure data aggregation scheme. The scheme modified the slicing and assembling technology by adopting a random distribution to decide the number of sliced data. The number of data pieces that each node slices its private data will not be a fixed number anymore and achieves better preservation of privacy and saves more energy for data aggregation. In [22], ESPART presents a novel energy-saving privacy-preserving aggregation scheme, which uses characteristic of the data aggregation tree structure to reduce communication overhead, assigns the random time pieces to nodes to avoid collision, and limits the scope of collision data to reinforce data loss resilience. Compared with the SMART, ESPART can preserve data privacy, get accurate data aggregation results while taking the same epoch duration as TAG, and have less communication overhead.

In ESPART model, a const $\text{MinDeg}$ is set. If the indegree of a node is less than $\text{MinDeg}$, data in this node needs to be sliced. When preserving privacy, it begins at the nodes whose indegree equals 1 in the WSN. These nodes slice one piece of data to their neighbor. And then do the same thing to the nodes whose indegree equals 2, till all the indegree
of nodes in the WSN is not less than MinDeg, and then the privacy-preserving process is ended. The process of mix and aggregation is the same as SMART model.

All the models based on SMART above send the slice to the neighbors randomly. In SMART scheme, a slice increases both the privacy-preserving efficacy of sending node and receiving node. But randomly slicing may lead the indegree of some nodes to getting large which is a redundancy of the privacy-preserving efficacy to the WSN.

The redundancy of the privacy-preserving efficiency means that the privacy-preservation efficacies of some nodes are far larger than other nodes. The redundancy costs more communication overhead which will shorten the lifetime of WSN.

In this paper, we present a balance privacy-preserving data aggregation model. The balance mechanism in the model ensures that slice can be sent to the nodes which have lower privacy preservation and enhances the privacy-preserving efficacy. At the same time, the model has less communication overhead and can prolong the lifetime of wireless sensor networks.

3. Preliminaries

In this section, we explain our network model, as well as our assumptions and the key pre-distribution scheme used in our model.

3.1. Network Model. Here, we consider a WSN network including \( N \) nodes and the network is connected. All the nodes build a graph \( G(V,E) \), where \( V \) is the set of the sensor nodes, \( E \) is the link of the nodes, and \( |V| = N \). In the proceeding of data aggregation, sensor nodes will be organized as a tree topology over \( G \) according to the typical protocol TAG [25].

Sensor nodes collect various data from monitoring environment and send them to the base station with suitable data aggregation schemes. In our model, we consider an additive aggregation function. It is a basic aggregation function because plenty of aggregation functions, such as count, average, and variance, can be deduced to the additive aggregation function. Data aggregation function is usually defined as follows:

\[
y(t) = f(d_1(t), d_2(t), \ldots, d_N(t)),
\]

where \( d_i(t) \) is the data which sensor node \( i \) gathered at time \( t \).

3.2. Key Distribution. To prevent attackers from eavesdropping, some messages are usually encrypted before sending the data. The following is the brief review of the random key distribution mechanism proposed in [25] which will be used in our model.

Firstly, a large key pool of \( K \) keys and their corresponding identities are generated. Each sensor node in WSN chooses \( k \) keys randomly from the key-pool and finds out which neighbors share a common key with itself by exchanging discovery messages. A secure link exists between two neighboring nodes only if they share a key. If two neighboring nodes cannot share a key but they can be connected by a link consisting of some nodes, this link can be the secure link between these two nodes.

In the random key distribution mechanism mentioned above, the probability that any pair of nodes possess at least one common key is

\[
P_{\text{connect}} = 1 - \frac{(K - k)^2}{(K - 2k!)K!},
\]

and the probability that any other node can overhear the encrypted message by a given key is

\[
P_{\text{overhear}} = \frac{k}{K}.
\]

Assume there are 10000 keys in the key pool, that is, \( K = 10000 \), and each node chooses 300 keys randomly, that is \( k = 300 \). The probability that any pair of nodes can find a shared key in common is \( P_{\text{connect}} = 99.9\% \) by (2). These pairs who do not share a common key can use the path-key establishment procedure which is described above to establish a shared key. Once a pair of nodes selects a shared key, the probability that any other node owns the same key is \( P_{\text{overhear}} = 0.3\% \), which is very small.

4. Balance Privacy-Preserving Data Aggregation Model

This section describes the details of the balance privacy-preserving data aggregation model (BPDA).

BPDA model considers the balance of the privacy-preserving efficacy in the whole WSN. In BPDA model, when nodes slice the data and send them to the neighbors, a balance mechanism is used to ensure that these slices will be sent to the nodes which have a low privacy preservation efficacy. This mechanism holds all the nodes at a similar privacy preservation efficacy, reduces the redundancy of the privacy preservation efficacy, and prolongs the lifetime of the WSN.

Figure 1 is an example to show the difference of balance slicing scheme and random slicing scheme. After building a tag tree, nodes 1 to 6 prepare to slice the data. In random slicing scheme (a), the minimum degree of all six nodes is 3 while degrees of node 3 and 5 reach 5, so the two nodes have privacy-preserving redundancy. So, we can adopt balance slicing scheme (b) which only increases the degree of those nodes whose degree is 3. So, only the degree of node 3 is 4, while others are 3. Less degree leads to fewer slices which can reduce communication overhead.

BPDA model consists of three phases as shown in Figure 2.

(1) Preparing Phase. An aggregation tree is constructed according to the standard aggregation protocol TAG. Each node records its own degree and computes the threshold of slice to prepare for data aggregation.
4.1. Preparing Phase. In the preparing phase, after establishing the TAG aggregation tree, each node records its own degree and then broadcasts the degree to its neighbors in one hop.

Next, each node prepared to utilize slicing and mixing technology in order to preserve data privacy. The number of slice plays an important role for the privacy-preservation which decides the minimum of the privacy preservation efficacy to the WSN. In the existing schemes, the number of slice is estimated according to administrator experience. In our BPDA scheme, we will give a principle to decide the slicing number.

In many applications, the network manager may expect that the exposed probability of the privacy data is not more than a const $Q$. A degree threshold $\text{MinDeg}$ is computed according to the experience probability. We assume that $P_i$ is the exposed probability of a data collected by node $i$, so we have

$$\max_i P_i \leq Q. \quad (4)$$

In this paper, the data collected by node $i$ is exposed only if all the messages both sent to and received from this node are exposed. Obviously, the sum of these messages equals the degree of this node, so we get

$$P_i = q^{d_i}, \quad (5)$$

where $d_i$ is the degree of node $i$, and $|d_i| \geq 1$ $q$ is the probability that one message is exposed. So $P_i$ is the increasing function of $q$. In reality, $q$ is bounded, so the formula (4) can be changed into

$$\max_i q^{d_{\text{max}}} \leq Q. \quad (6)$$

According to the formula (6), we can get the minimum degree which satisfies the formula (6) as follows:

$$d_{\text{min}} = \log_{q_{\text{max}}} Q. \quad (7)$$

On the other hand, $d_{\text{min}}$ should be a integer, so formula (7) can be adjust to

$$d_{\text{min}} = \lceil \log_{q_{\text{max}}} Q \rceil + 1, \quad (8)$$

where $\lceil \log_{q_{\text{max}}} Q \rceil$ is rounded down of $\log_{q_{\text{max}}} Q$.

In BPDA model, the threshold of the degree can be set according to the process above, so we have

$$\text{MinDeg} = \lceil \log_{q_{\text{max}}} Q \rceil + 1. \quad (9)$$

Figure 3 shows the relationship of the degree threshold MinDeg and the exposed probability $q$. Given $Q = 0.0001$, that is, the probability that all nodes in networks are exposed, is less than 0.01%, the value of MinDeg increases with the increasing of probability $q$. When the value of $q$ is less than 0.02, the value of MinDeg is 2. While $q$ increases to 0.06, MinDeg increases to 4.

4.2. Privacy-Preserving Phase. After the preparing phase, the base station computes the whole time of privacy-preserving phase $T_p$ by estimating the slicing time $t_s$ in one round combining with the threshold MinDeg, which satisfies the condition as follows:

$$T_p = \text{MinDeg} \times t_s. \quad (10)$$
4.2.1. Energy Based. In energy based algorithm, a threshold $\overline{E}_{tr}$ is set as the average of remaining energy in the neighbors of node $i$. Meanwhile, $t_w$ is the waiting time and $E_{ri}$ means the remaining energy of node $i$.

Firstly, a receiving node should be determined by the connected probability. The connected probability in this part is as follows:

$$p_{i,k} = \frac{1/N_j}{\sum_{j \in \text{neighbours}_k} 1/N_j},$$

(14)

where $p_{i,k}$ is the probability that node $k$ connects to its neighbor $i$. $N_j$ is the remaining energy of node $i$. neighbours$_k$ is the neighbor set of node $k$.

Secondly, if $E_{ri} \geq \overline{E}_{tr}$, node $i$ sends one slice to the receiving node. If $E_{ri} < \overline{E}_{tr}$, node $i$ waits for a $t_w$ time. In $t_w$, if node $i$ receives no slice, it will send one slice to the receiving node.

This algorithm balances the energy consumption and prolongs the lifetime in the WSN, but it may cause some redundancy of the privacy preservation efficacy. The model using this algorithm is called E-BPDA model.

4.2.2. Degree Based. In degree based algorithm, only one rule is considered; that is, a node with higher degree has lower probability to connect and to be connected. The connected probability is as follows:

$$p_{i,k} = \frac{1/d_i}{\sum_{j \in \text{neighbours}_k} 1/d_j},$$

(15)

where $p_{i,k}$ is the probability that node $k$ connects to its neighbor $i$. $d_i$ is the degree of node $i$. neighbours$_k$ is the neighbors set of node $k$.

This algorithm reduces the redundancy of the privacy preservation and balances the privacy preservation of the whole WSN, but it may cause a little unbalance of the energy consumption. The model using this algorithm is called D-BPDA model.

4.2.3. Both Energy and Degree Based. Energy based algorithm and degree based algorithm are complementary to each other. So the cooperation of these two types is considered. Firstly, similarly to energy based algorithm, $E_{ri}$, $t_w$, and $\overline{E}_{tr}$ should be computed or set.

Firstly, a receiving node should be determined by the connected probability. The connected probability in this part is as follows:

$$p_{i,k} = \frac{1/(d_i \cdot N_j)}{\sum_{j \in \text{neighbours}_k} (1/(d_j \cdot N_j))},$$

(16)

where $p_{i,k}$ is the probability that node $k$ connects to its neighbor $i$. $N_j$ is the remaining energy of node $i$. $d_i$ is the degree of node $i$. neighbours$_k$ is the neighbor set of node $k$.

Secondly, if $E_{ri} \geq \overline{E}_{tr}$, node $i$ sends one slice to the receiving node. If $E_{ri} < \overline{E}_{tr}$, node $i$ waits for a $t_w$ time. By the
Base station
Estimate the time $t_s$ that a round of slicing in whole WSN
Compute the whole slicing time $T_p$, s.t.

$$T_p = \text{MinDeg} \times t_s$$

Broadcast the values MinDeg, $T_p$ and $t_s$ to the whole sensor nodes

Sensor node $i$
Receive the values MinDeg, $T_p$ and $t_s$ from the base station
degree$_i$ is the degree of node $i$
data$_i$ is the data of node $i$

In one round of slicing time
If degree$_i < \text{MinDeg}$
   Find the neighbors of node $i$
   Compute the connected probability of each neighbor
   Confirm the receiving node (assume that is node $i$)
   Produce a slice slice$_{ij}$ in node $i$
   Send the slice$_{ij}$ to node $j$
   Update the data$_i$, s.t.
   data$_i$ = data$_i$ − slice$_{ij}$
   Update the degree$_i$, s.t.
   degree$_i$ = degree$_i$ + 1
   If receive a slice$_{ki}$ from node $k$
      Receive slice$_{ki}$
      Update the degree$_i$, s.t.
      degree$_i$ = degree$_i$ + 1
   End If
Else
   If receive a slice$_{ki}$ from node $k$
      Receive slice$_{ki}$
      Update the degree$_i$, s.t.
      degree$_i$ = degree$_i$ + 1
   End If
End If

Algorithm 1: Algorithm of slicing.

end of time $t_w$, if node $i$ receives no slice, it sends one slice to the receiving node.

This algorithm reduces the redundancy of the privacy preservation efficacy and balances the energy consumption at the same time. The model using this algorithm is called C-BPDA model.

4.3. Data Aggregation Phase. In this phase, each node sends its data to the base station along the aggregation tree.

5. Analysis of Privacy Preservation Efficacy

An evaluation method is necessary to compare different privacy-preserving schemes. One of such methods is proposed in [13] and is used by many other papers which can be described as follows.

Firstly, it assumes that $P_{\text{overhear}}$ is the probability that any node is eavesdropped. And the probability that any two nodes collude is $P_{\text{collude}}$. Moreover, the probability that these two probabilities are equal to each other is assumed in this method. So the formula is as follows:

$$P_{\text{overhear}} = P_{\text{collude}} = q. \quad (17)$$

Then, the probability that the private data of node $s$ is exposed for a given $q$ under either condition above in SMART algorithm is as follows:

$$P(q) = q^{J-1} \sum_{k=1}^{d_{\text{max}}} P(\text{in\_degree} = k) \cdot q^k, \quad (18)$$

where $J$ is the number of the slices, $d_{\text{max}}$ is the maximum of the indegree in the WSN, and $P(\text{in\_degree} = k)$ is the probability that the indegree of the node equals $k$.

Obviously, $J - 1$ is the outdegree of the node. So, $P(q)$ can be expressed generally as follows:

$$P(q) = \sum_{k=d_{\text{min}}}^{d_{\text{max}}} P(\text{degree} = k) \cdot q^k. \quad (19)$$

Actually, this evaluation considers the privacy preservation of the whole network instead of some certain node. As shown in Table 1, there are two networks, NW1 and NW2. Each of them has 8 nodes. Nodes in NW1 are not of the same degree, but in NW2 every node has the same degree of 3.
Table 1: The degree distributions of two networks.

| Nodes | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|-------|----|----|----|----|----|----|----|----|
| Degree | NW1 | 4  | 6  | 2  | 6  | 6  | 4  | 6  |
| NW2   | 3  | 3  | 3  | 3  | 3  | 3  | 3  | 3  |

According to the Table 1 and formula (18), we have

\[ P_{NW1}(q) = \sum_{k=d_{min}}^{d_{max}} P(\text{degree} = k) \cdot q^k = \frac{1}{8}q^2 + \frac{3}{8}q^4 + \frac{1}{2}q^6, \]

\[ P_{NW2}(q) = \sum_{k=d_{min}}^{d_{max}} P(\text{degree} = k) \cdot q^k = q^3. \]

If \( q = 0.2 \), we get \( P_{NW1} = 0.0056 \) and \( P_{NW2} = 0.008 \). Obviously, NW1 is more robust than NW2 as \( P_{NW1} < P_{NW2} \). However, as shown in Table 1, there is a node whose degree is 2 in NW1 which is the most easily to be disclosed both in NW1 and NW2. So this evaluation method describes the global privacy preservation efficacy instead of focusing on a specific node’s privacy preservation which is of more concern in practical application.

6. Simulation

In this section, a wireless sensor network with 800 nodes is considered, and these nodes are randomly deployed over 400 x 400 area. The energy of each node is 0.5 J. We apply TAG scheme [25] which is a typical data aggregation scheme in the simulation. We study the performances of BPDA model in four aspects with simulation which are degree distribution, privacy preservation efficacy, communication overhead, and lifetime. BPDA models will be compared with SMART model and ESPART model in these performances.

6.1. Degree Distribution. In this section, a node with degree of 2 is regarded as privacy-preserved enough.

Figure 4 shows degree distribution in different models. In TAG, the minimum degree is 1 and the maximum degree is 9 while 80 percent of the other models in network take the minimum degree. After privacy preserving, all the schemes increase the minimum degree to 2 and the maximum degree is increased too. In three BPDA models, the D-BPDA and C-BPDA only increase the maximum degree from 9 to 10. In the E-BPDA model, the maximum degree increases to 11 which is the same in the ESPART model. Meanwhile, the SMART model increases the maximum degree to 16. The increasing of maximum degree means that some nodes which need not to be privacy preserved are preserved. This is a main reason that causes the redundancy of the privacy preservation.

6.2. Privacy Preservation Efficacy. Here, the evaluation method of the privacy preservation efficacy in [22] is adopted. Figure 5 shows the exposed probability of nodes in different models. In Figure 5, the exposed probability of nodes in BPDA models is higher than that of SMART and ESPART models because this evaluation method works from a global view of the whole WSN. In many cases, the larger the sum of degrees is, the lower exposed probability the model has. It seems that SMART and ESPART models have more ability on privacy preservation because they pay much more on the redundancy when some nodes have high privacy preservation with rather high degree after the operation. The BPDA models consider the redundancy problem and put the algorithms only effect on the nodes with minimum degree. Although their exposed probabilities are higher than others, they are still kept in the similar level.

6.3. Communication Overhead. As to the communication overhead, the amount of the sending data in slicing step is considered. Tables 2 and 3 show the amount of sending data of different schemes SMART, ESPART, D-BPDA, E-BPDA, C-BPDA, and their percentage to SMART at conditions \( J = 2 \), \( \text{MinDeg} = 2 \) and \( J = 3 \), \( \text{MinDeg} = 3 \).

On the other hand, the degree of more than 50 percent of nodes in three BPDA models is 2 while ESPART and SMART schemes increase more nodes’ degrees which are redundancy.

In three BPDA models, E-BPDA considers so many energy balances that its redundancy is the most. D-BPDA has the less redundancy by considering how to reduce it. The C-BPDA which combines both D-BPDA and E-BPDA leads less redundancy than the E-BPDA.
Figure 5: The exposed probability of nodes in different algorithms.

Table 2: The communication overhead and percentage to SMART, at $J = 2$, MinDeg = 2.

| Schemes | Sending data | Percentage to SMART |
|---------|--------------|---------------------|
| SMART   | 800          |                     |
| ESPART  | 621          | 77.63%              |
| D-BPDA  | 440          | 55%                 |
| E-BPDA  | 494          | 61.75%              |
| C-BPDA  | 429          | 53.63%              |

Table 3: The communication overhead and percentage to SMART, at $J = 3$, MinDeg = 3.

| Schemes | $J = 3$, MinDeg = 3 | Percentage to SMART |
|---------|---------------------|---------------------|
| SMART   | 1600                |                     |
| ESPART  | 924                 | 57.75%              |
| D-BPDA  | 805                 | 50.31%              |
| E-BPDA  | 896                 | 56%                 |
| C-BPDA  | 800                 | 50%                 |

So the BPDA models reduce the communication overhead obviously.

When $J = 2$, comparing with SMART scheme which has the largest amount of sending data 800, ESPART scheme is 621 and 77.63% to the SMART scheme, D-BPDA is 55%, E-BPDA is 61.75%, and C-BPDA is 53.63%.

Similar to $J = 2$, when $J = 3$, the communication overhead of SMART is still the largest, and it reaches 1600 which is twice to the value of $J = 2$. ESPART scheme is 577.75% to the SMART scheme, D-BPDA is 50.31%, E-BPDA is 56%, and C-BPDA is 50%.

From the data of Tables 2 and 3, we can see that he BPDA models send less data which means they have a lower communication overhead. The ESPART model and BPDA models are closer in sending data with the increasing $J$. As a general rule, the BPDA models can preserve the data privacy well while using slice with $J = 2$ or $J = 3$ as well as reducing the communication overhead.

In three BPDA models, the E-BPDA model considers more of the energy balance of the whole network, so it causes the higher communication overhead. Other two models both consider the degree balance which causes lower communication overhead. And in D-BPDA and C-BPDA, the communication overheads are at the same level.

6.4. Lifetime. In the simulation of lifetime, we assume that all sensor nodes have an initial energy which is 0.5 J. The data packet size is 1000 bits. The minimum degree in the network is 2 after privacy preservation. A WSN cannot operate when more than 20% of the sensor nodes are out of work. And the number of nodes in network is 800. Therefore, the network lifetime is defined as the time when 160 sensor nodes are discharged.

Nodes consume energy both in sending and receiving data according to [26, 27]. In this paper, we use the model that the pass loss exponent is 2. The model is as follows.

A $k$-bit data packet is transmitted and the energy consumption of sending node is given by

$$ E_s = \epsilon_1 \times k + \epsilon_2 \times d^2 \times k, $$

(21)

where $d$ is the distance between the two sensor nodes and $\epsilon_1 = 50 \text{ nJ/bit}$, $\epsilon_2 = 100 \text{ pJ/bit/m}^2$.

A $k$-bit data packet is transmitted, and the energy consumption of receiving node is given by

$$ E_r = \epsilon_1 \times k. $$

(22)

Figure 6 shows that the drained nodes appear in about 700th round in the three BPDA schemes. And there is a bifurcation point at about the 900th round. Before the demarcation point, the increasing of the drained nodes in all models is at the same trace. After the point, the drained nodes in D-BPDA models increase to 160 in about 200 rounds, and then the lifetime is up. In the E-BPDA and C-BPDA models, the increasing of drained nodes is slower than D-BPDA model. So it can prolong the lifetime when considering the energy balance which can balance the energy consumption of the network.

7. Conclusion

In wireless sensor networks, sensitive information that sensor nodes gathered is prone to be leaked for the hostile environment. Privacy-preserving has become an important issue in data aggregation. A balance privacy-preserving data aggregation model based on slicing and mixing technology is proposed in this paper. Firstly, a degree threshold is computed according to security requirement of the WSN. Compare with fixed or random slicing, the proposed slicing method emphasizes that sensor node sends the slices to its neighbors refer to the degree threshold and ensures that the slices can be sent to the nodes which have lower privacy preservation. So, it reduces the redundancy and increases the
privacy-preservation efficacy. Furthermore, according to the influence factor in real application, energy based E-BPDA, degree based D-BPDA, and both energy and degree C-BPDA three different schemes are presented to keep the privacy-preserving data aggregation balance. Simulation shows that E-BPDA model has a longer lifetime, D-BPDA model has a lower communication overhead, and the C-BPDA combines the advantage of the E-BPDA and D-BPDA.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant 61201159, the Beijing Natural Science Foundation under Grant (4132057), and the Beijing Municipal Education Commission on Projects (SQKM201510016013).

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” IEEE Communications Magazine, vol. 40, no. 8, pp. 102–114, 2002.
[2] D. Culler, D. Estrin, and M. Srivastava, “Overview of sensor networks,” IEEE Computer, vol. 37, no. 8, pp. 41–49, 2004.
[3] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, “Wireless sensor networks for habitat monitoring,” in Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications (WSNA’02), pp. 88–97, Atlanta, Ga, USA, September 2002.
[4] N. Xu, S. Rangwala, K. K. Chintalapudi et al., “A wireless sensor network for structural monitoring,” in Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys ’04), pp. 13–24, ACM, November 2004.
[5] B. Krishnamachari, D. Estrin, and S. Wicker, “The impact of data aggregation in wireless sensor networks,” in Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops, pp. 575–578, IEEE, Vienna, Austria, 2002.
[6] J. Yick, B. Mukherjee, and D. Ghosal, “Wireless sensor network survey,” Computer Networks, vol. 52, no. 12, pp. 2292–2330, 2008.
[7] H. Cam, S. Ozdemir, P. Nair, and D. Muthuvinashippan, “ESPDA: energy-efficient and secure pattern-based data aggregation for wireless sensor networks,” in Proceedings of the IEEE Conference on Sensors, pp. 732–736, IEEE, Piscataway, NJ, USA, October 2003.
[8] L. Hu and D. Evans, “Secure aggregation for wireless networks,” in Proceedings of the Symposium on Applications and the Internet Workshops, pp. 384–391, IEEE Press, 2003.
[9] B. Przydatek, D. Song, and A. Perrig, “SIA: secure information aggregation in sensor networks,” in Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys ’03), pp. 255–265, Los Angeles, Calif, USA, November 2003.
[10] S. Ozdemir and Y. Xiao, “Secure data aggregation in wireless sensor networks: a comprehensive overview,” Computer Networks, vol. 53, no. 12, pp. 2022–2037, 2009.
[11] H. Alzaid, E. Foo, and J. G. Nieto, “Secure data aggregation in wireless sensor network: a survey,” in Proceedings of the 6th Australasian Information Security Conference (AISC ’08), pp. 93–105, January 2008.
[12] H. O. Sanli, S. Ozdemir, and H. Cam, “SRDA: secure reference-based data aggregation protocol for wireless sensor networks,” in Proceedings of the IEEE VTC Fall Conference, pp. 4650–4654, Los Angeles, Calif, USA, September 2004.
[13] N. Li, N. Zhang, S. K. Das, and B. Thuraisingham, “Privacy preservation in wireless sensor networks: a state-of-the-art survey,” Ad Hoc Networks, vol. 7, no. 8, pp. 1501–1514, 2009.
[14] R. Bista and J.-W. Chang, “Privacy-preserving data aggregation protocols for wireless sensor networks: a survey,” Sensors, vol. 10, no. 5, pp. 4577–4601, 2010.
[15] C. Castelluccia, E. Mykletun, and G. Tsudik, “Efficient aggregation of encrypted data in wireless sensor networks,” in Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, pp. 109–117, July 2005.
[16] J. Girao, D. Westhoff, and M. Schneider, “CDA: concealed data aggregation for reverse multicast traffic in wireless sensor networks,” in Proceedings of the IEEE International Conference on Communications (ICC ’05), vol. 5, pp. 3044–3049, IEEE, May 2005.
[17] S. Ozdemir and Y. Xiao, “Integrity protecting hierarchical concealed data aggregation for wireless sensor networks,” Computer Networks, vol. 55, no. 8, pp. 1735–1746, 2011.
[18] W. B. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher, “PDA: privacy-preserving data aggregation in wireless sensor networks,” in Proceedings of the 26th IEEE International Conference on Computer Communications (IEEE INFOCOM ’07), pp. 2045–2053, Anchorage, Alaska, USA, May 2007.
[19] W.-B. He, N. Hoang, X. Liu, K. Nahrstedt, and T. Abdelzaher, “iPDA: an integrity-protecting private data aggregation scheme for wireless sensor networks,” in Proceedings of the IEEE Military Communications Conference (MILCOM ’08), pp. 1–7, San Diego, Calif, USA, November 2008.
[20] M. M. Groat, W. Hey, and S. Forrest, “KIPDA: K-indistinguishable privacy-preserving data aggregation in wireless sensor networks,” in Proceedings of the 30th IEEE International Conference on Computer Communications (INFOCOM ’11), pp. 2024–2032, April 2011.

[21] 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.

[22] 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.

[23] G. Yang, S. Li, X. Xu, H. Dai, and Z. Yang, “Precision-enhanced and encryption-mixed privacy-preserving data aggregation in wireless sensor networks,” International Journal of Distributed Sensor Networks, vol. 2013, Article ID 427275, 12 pages, 2013.

[24] S. Madden, M. J. Franklin, and J. M. Hellerstein, “TAG: a Tiny AGgregation service for Ad-hoc sensor networks,” in Proceedings of the 5th Symposium on Operating Systems Design and Implementation (OSDI ’02), Boston, Mass, USA, December 2002.

[25] L. Eschenauer and V. D. Gligor, “A key-management scheme for distributed sensor networks,” in Proceedings of the 9th ACM Conference on Computer and Communications Security, pp. 41–47, November 2002.

[26] M. Hussaini, H. Bello-Salau, A. F. Salami, F. Anwar, A. H. Abdalla, and M. R. Islam, “Enhanced clustering routing protocol for power-efficient gathering in wireless sensor network,” International Journal of Communication Networks and Information Security, vol. 4, no. 1, pp. 18–28, 2012.

[27] W. Qin, M. Hempstead, and Y. Woodward, “A realistic power consumption model for wireless sensor network devices,” in Proceedings of the 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks, pp. 286–295, September 2006.