Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 134509, 9 pages
http://dx.doi.org/10.1155/2013/134509

Research Article

ESMART: Energy-Efficient Slice-Mix-Aggregate for Wireless Sensor Network

Chaoran Li and Yun Liu

Key Laboratory of Communication & Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Yun Liu; liuyun@bjtu.edu.cn

Received 30 July 2013; Accepted 19 November 2013

Academic Editor: Ying-Hong Wang

Wireless sensor network consists of a large number of resource-constrained sensor nodes and is usually deployed in unattended area to collect specific information. Energy consumption is always a major concern in the research field of wireless sensor network. Thus, data aggregation schemes emerged and were deployed for prolonging network lifetime by reducing data transmitted within the network. Meanwhile, along with the wide application of the data aggregation schemes, the security issues of data aggregation have been increasingly drawing attention and the designing of secure data aggregation scheme is becoming a hot spot. In this paper, we proposed an energy-efficient secure data aggregation scheme, ESMART: energy-efficient slice-mix-aggregate based on the technique of data slicing and mixing. The proposed scheme performs secure data aggregation in a more efficient way while keeping a good performance of privacy preservation. And the simulation result shows that the security performance of ESMART scheme is better than that of some existing and widely used schemes.

1. Introduction

A wireless sensor network is composed of a large number of resource-constrained sensor nodes. The base station of the network sends queries to the child nodes, and the sensor nodes that received the queries will collect the information following the request of the query. The data acquired by sensor nodes will be transmitted to the base station with multihop transmission scheme. Obviously, the energy consumption of transmission will affect the life cycle of the sensor nodes dramatically. Hence, the technology of data aggregation has been widely used in wireless sensor network as an effective mechanism to reduce the amount of data transfer in the network. In the mechanism of data aggregation, the nodes which aggregate raw data from their neighbor nodes and process it by specific algorithms are called aggregators, and the other sensor nodes which center on the aggregators and only perform information collecting and transmitting are called leaf nodes. The processed and aggregated data will be returned to the base station through the aggregation tree constructed dynamically. After the application of data aggregation technology, the amount of data transfer in WSNs has been reduced greatly and then prolonged the operating life of the network.

Meanwhile, some security problems have arisen because of the deployment of data aggregation scheme. As mentioned above, sensor nodes are usually resource constrained, the computing and communicating capacity of the nodes are very limited which makes the designing of the secure scheme in WSN more difficult than in traditional wired network, and most of the existing, mature secure protocols could not be applied to WSN directly. From the perspective of the adversary, the sensor nodes are usually deployed in unattended even hostile environment which makes sensor nodes very vulnerable to the attacks. We take a simple and widely used data aggregation scheme TAG: Tiny Aggregation proposed by Madden et al. in [1] as an example. We can see that there is no protection of data privacy in this scheme which means a captured node in the network can eavesdrop on and decrypt the raw readings of neighbor nodes easily without being detected and then lead to a serious security problem which is not acceptable [2, 3]. On the basis of the above analysis, a good secure data aggregation scheme for WSN should be energy efficient while ensuring data security. In this paper, we propose a novel secure data aggregation scheme for WSN, ESMART (energy-efficient slice-mix-aggregate), which addresses both energy-efficient and privacy preservation of the network. To achieve the privacy preservation of the network, we introduced data slicing and mixing technique into the scheme designing. A secure data aggregation scheme based on data slicing and mixing technique was detailed in [4]; in this scheme, in order to achieve the preservation of data privacy in the network, the raw readings of the sensor nodes will be sliced into pieces and part of the pieces will be encrypted and sent to the neighbor nodes. But this comes at the price of some additional communication cost which is shown in the simulation result of SMART scheme; the communication cost of SMART is about J times the communication cost of TAG, where J denotes the number of data pieces. Thus we introduced the dynamic J value into the scheme which can effectively decrease the communication overhead while maintaining the privacy preservation of the network. Meanwhile, a lower communication overhead means a lower transmission collision probability, which leads to a higher aggregation accuracy in the network. Moreover, in ESMART, the times of data slicing of a sensor node are only known by itself, so even if a malicious node eavesdropped on several data pieces from the node, it is not sure that all the data pieces of the node are eavesdropped on. Thus, it is more difficult for the malicious nodes to steal private data. Given the above, the use of data slicing and mixing technique in ESMART is more efficient which consequently reduces the communication overhead and increases the aggregation accuracy effectively.

Wireless sensor network is usually deployed in unattended environment and consisted of a large number of low-powered, resource-constrained sensor nodes. Each node in the network can sense the environment information as requested by the base station and perform wireless communication within a small range of its location. To reduce the energy consumption of the nodes during the transmission, data aggregation scheme has been widely used in WSN. Madden et al. proposed TAG scheme as a simple data aggregation scheme based on the assumption that every node in the network is trusted, and there is no protection of data privacy in this scheme. But in practice the application of WSN is facing various kinds of security threats. A good secure data aggregation scheme should possess the capability to resist the security threats while staying within energy constraint. To meet this requirement, some secure data aggregation schemes have been provided. He et al. proposed a novel data aggregation scheme with disjoint aggregation tree structure to ensure data integrity in [5]. Liu et al. proposed a secure data aggregation scheme HEEP with a novel formation method of aggregation tree in [6] in 2013. Li et al. proposed a secure data aggregation scheme based on the improved SMART scheme which is more efficient [7]. Both Li et al. [8] and Bista and Chang [9] proposed survey papers of the privacy-preserving techniques for WSN. Meanwhile, some researchers analyzed the existing network attack technologies against WSN and proposed corresponding solutions [10–12]. Zhu et al. proposed a secure data aggregation scheme, in which the base station composes a secret configuration matrix and each sensor node is preloaded with a limited part of the matrix known as a secret share containing certain local instructions [13]. Dong and Li presented a secure data aggregation approach based on monitoring in WSN [14]. In this paper, a grid-based network architecture for monitoring in WSN is designed and the algorithms for network division, initialization, and grid tree construction are proposed. Lei et al. proposed a secure data aggregation solution based on dynamic multiple cluster key management model in [15]; this key management model consumes little storage space and can resist the collusion attack effectively. Sun et al. proposed a lightweight secure data aggregation protocol to find the compromised nodes and help the base station to verify the final results [16]. Poornima and Amberker proposed a secure data aggregation scheme in [17]. This scheme provides end-to-end data privacy and can effectively reduce the energy consumption.

2. Network Model and the Requirement of Design

In this paper, we used aggregation tree to organize the sensor nodes in the network and consider three types of nodes in the network, base station, aggregator node, and leaf node. There is only one base station in the network which can be seen as the network control center with unlimited resources and the final destination for the aggregation results. As the root of the aggregation tree, it is responsible for broadcasting queries to all the other nodes and receiving and processing the aggregation results. A sensor node in the network can choose to be an aggregator node with the probability P_c which is a preset value. The aggregator nodes are responsible for aggregating target information and the query results collected by other nodes. The other nodes in the network will become leaf nodes and will be in charge of collecting and transmitting information to their neighbor aggregator nodes. Each node in the network can only communicate with its neighbor nodes within h hops; for a dense WSN, we take $h = 1$. It is important to note that the proposed aggregation scheme is used to perform addictive aggregation functions which are not restrictive like sum, average, variance, and so forth. And the concrete details of addictive aggregation algorithms are not covered in this paper.

2.1. Network Model. We consider a wireless network consisting of N nodes and one base station. Each node in the network has the functionalities of sensing, computing, and transmitting. The collected primitive data of node i can be expressed as R_i, and the aggregation function is defined as $A_i = f(R_i)$, where A_i denotes the aggregated data of node i. Meanwhile, we introduced time variable into the function and obtained $A_i(t) = f(R_i(t))$, where $R_i(t)$ denotes the raw data collected by node i at time t and $A_i(t)$ means the aggregation results of $R_i(t)$. To prevent the private data from being eavesdropped on in the link level, we used random key distribution mechanism in the scheme whose work mechanism was introduced by Eschenauer and Gligor in [18], and there are also some other good key management schemes that have been proposed which are not covered in this paper [19–21]. In the phase of key distribution, a large key pool is generated and part of the keys in the pool will be drawn to form a key ring for a node. After the key distribution, any nodes that find out the neighbor which shared common keys with them can maintain a secure communication link with the neighbor. In this paper, we consider K_p is the size of key pool, and K_r keys are taken out from the key pool randomly. Then we can express the probability that any third party has the shared key of a secure link as below:

$$P_{\text{eavesdrop}} = \frac{K_r}{K_p}. \tag{1}$$

The formula above can also be interpreted as the probability that a secure link between two nodes is eavesdropped on by a third node and we can see from the formula that the larger the value of K_p is, the smaller $P_{\text{eavesdrop}}$ is, the more secure the link is. If the size of key pool is big enough, this key distribution mechanism can ensure that the value of $P_{\text{eavesdrop}}$ in the network is around 0.1%, under the condition that any pair of nodes in the network shares at least one common key with each other.

2.2. The Design Goals of the Scheme

Efficiency. Due to the energy constrained character of sensor nodes, data aggregation scheme was applied to reduce the communication overhead, but consequently the application of secure scheme for data aggregation would inevitably cause some security problems. Therefore, how to maintain a balance between security and energy consumption has become one of the core issues in the designing of secure data aggregation scheme. In this paper, the proposed scheme improved the network security while keeping the network system at a low energy cost level.

Privacy Preservation. The preservation of private data is always a critical security issue in the design of secure data aggregation scheme. It means any node in the network could only know the private data of itself. Some of the widely used data aggregation schemes like TAG ignored this issue which lead to the fact that a malicious node can eavesdrops on the communications of its neighbors to steal the private data easily and then cause a major security problem. Thus, the preservation of data privacy is a key point in the designing of the proposed scheme in this paper.

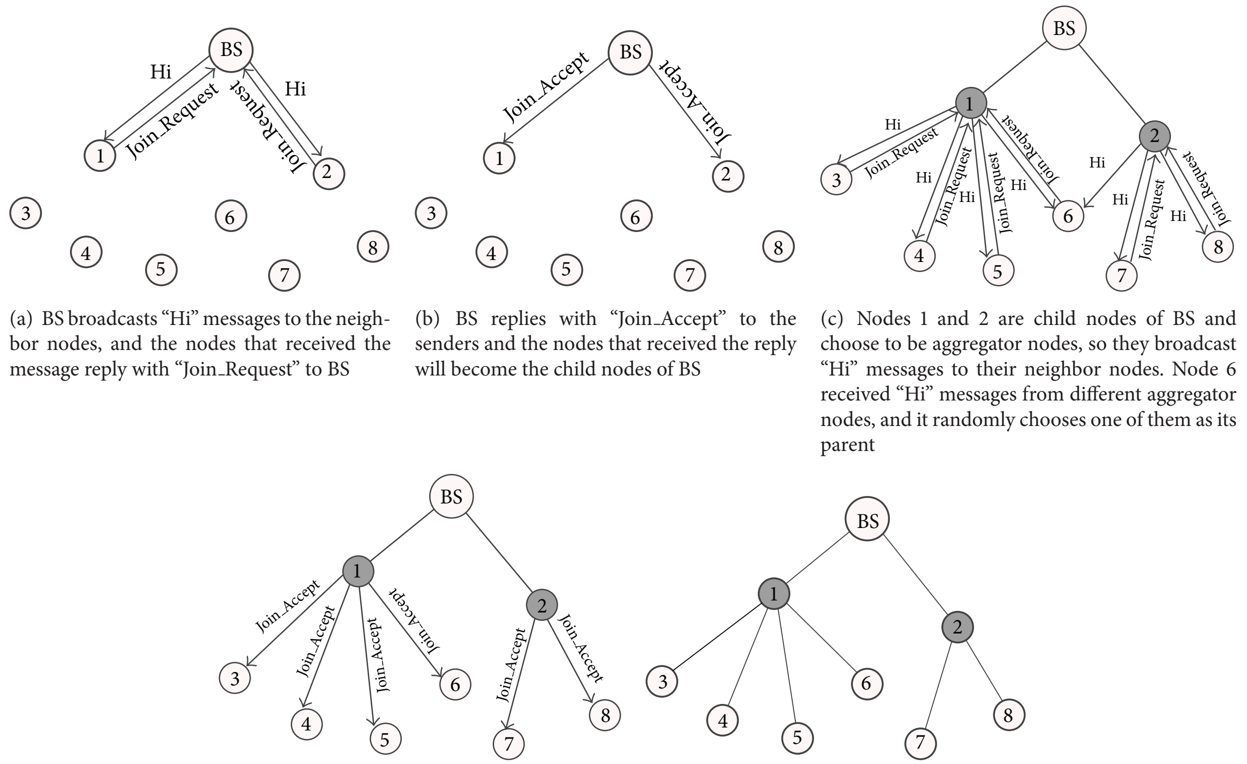
Accuracy. The accuracy of the aggregation result is a crucial criterion in assessing the performance of data aggregation scheme. It is affected by the success rate of data transmission in the network directly, and the factors influencing the success rate of data transmission include data loss, transmitting collision, and so forth. In a certain time, the more the data is transmitted in the network, the higher the risk of data loss or transmitting collision. Therefore, in the designing of the proposed scheme, we need to reduce the data traffic in network while providing the preservation of data privacy. In other words, the more efficient the scheme is, the more accurate the aggregation result would be.

3. The Proposed Secure Data Aggregation Scheme

3.1. The Formation of Aggregation Tree. In the proposed scheme, an aggregation tree rooted at the base station needs to be formed for organizing the sensor nodes in the network. First of all, the base station broadcasts "Hi" message to all the neighbor nodes within one hop. Any nodes without parent that received "Hi" message should reply with "Join_Request" message to the originator, but if the node received multiple "Hi" messages from different senders, then it should randomly select one of them to be its parent node. Once a message of "Join_Request" is received, the BS accepts the node to be its child node by replying to the message of "Join_Accept." The probability that a child node becomes an aggregator node is P_a which is a preset value, and the rest of the nodes will become leaf nodes. That means there are $N \cdot P_a$ aggregator nodes and $N \cdot (1 - P_a)$ leaf nodes in the network consisting of N sensor nodes. The aggregator nodes will continue broadcasting "Hi" messages to find their child nodes. If a node did not receive "Hi" message, it will broadcast "No_Parent" message to its neighbor nodes to find its parent node, and the aggregator node that received this message will accept it as its child node. After the formation of the aggregation tree, if any aggregator node in the aggregation tree were to fail, the parent of the failed aggregator node would broadcast "Hi" messages and try to communicate with the child nodes of the failed aggregator node. After receiving "Hi" message, the child nodes of the failed aggregator node will reply "Join_Request" messages to the sender. The parent of the failed aggregator node will randomly select one of these nodes to be the new aggregator node and reject all the other join requests. Next, the new aggregator node will then begin to broadcast "Hi" messages and accept all the other nodes which do not have a parent as their child nodes according to the rules introduced above. The causes of node failure are numerous, such as runout of power and physical damage, compromised by a cyber-attack. In future work, we will propose a trust management system of data aggregation, for monitoring node behavior, and then analyze the causes of node failure to detect the attacking strategy of the attacker early. Figure 1 describes the whole procedure of formatting an aggregation tree.

3.2. The Overview of the Proposed Secure Data Aggregation Scheme. In this section, we present the work mechanism of ESMART that can be divided into three stages: data slicing, data mixing, and data aggregation. The detailed introduction is as follows.

Figure 2 describes an aggregation tree consisting of one base station, two aggregator nodes, and six leaf nodes. We set $J = 4$, which is a preset value that defines the maximum of j, which means the range of j is from 2 to J and obeys the probability distribution formula $F(J, j)$. The calculating concrete process of $F(J, j)$ will be introduced in Section 4. We describe the work mechanism of ESMART in this network environment. After the collecting of target data, a leaf node in the network slices the collected data



(a) BS broadcasts "Hi" messages to the neighbor nodes, and the nodes that received the message reply with "Join_Request" to BS

(b) BS replies with "Join_Accept" to the senders and the nodes that received the reply will become the child nodes of BS

(c) Nodes 1 and 2 are child nodes of BS and choose to be aggregator nodes, so they broadcast "Hi" messages to their neighbor nodes. Node 6 received "Hi" messages from different aggregator nodes, and it randomly chooses one of them as its parent

(d) Nodes 1 and 2 reply with "Join_Accept" to the child nodes

(e) The aggregation tree is formed

Figure 1: The formation of aggregation tree.

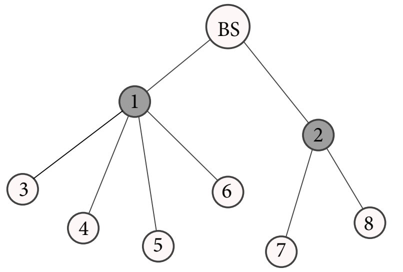


Figure 2: Aggregation tree ($h = 1$, $J = 4$).

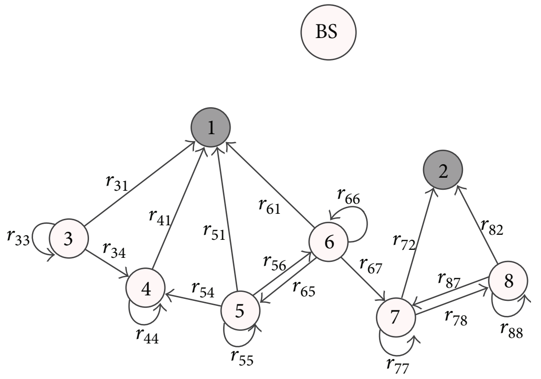


Figure 3: Data slicing.

into j pieces and randomly chooses $j - 1$ neighbor nodes which shared common keys with it to be the receivers of the data pieces and then sends $j - 1$ encrypted data pieces to the receivers and keeps the remaining one by itself. It is worth nothing that j is a variable which is different from SMART. When the node is an aggregator node, it just receives the data pieces sent by the leaf nodes instead of slicing the collected data of its own because the raw data collected by the aggregator node will be assembled with the received data pieces in the step of data mixing which achieved the target of hiding private data. Figure 3 describes the step of data slicing; r_a is the raw data collected by node a, and r_{ab} denotes a piece of data sent from node a to node b. We take node 3 and node 1 as an example; as a leaf node, node 3 sliced its primitive data into 3 pieces, r_{33}, r_{31}, and r_{34} and then sent the encrypted r_{31} and r_{34} to node 1 and node 4 separately. r_{33} is kept by itself. As an aggregator node, node 1 did not participate in the data slicing and just received the data pieces sent by leaf nodes.

After the stage of data slicing, the nodes decrypt all the received data pieces and sum them up with their raw data piece. The process can be described as $A_b = \sum_{a=1}^{N} r_{ab}$, $r_{ab} = 0$, if node b is not one of the receivers of node a, where A_b denotes the aggregation result of node b. And then we can express the final aggregation result as $R = \sum_{b=1}^{N} \sum_{a=1}^{N} r_{ab}$. Figure 4 described the step of data mixing; we take node 4 as an example: it sums up the received data pieces r_{34} and r_{54}

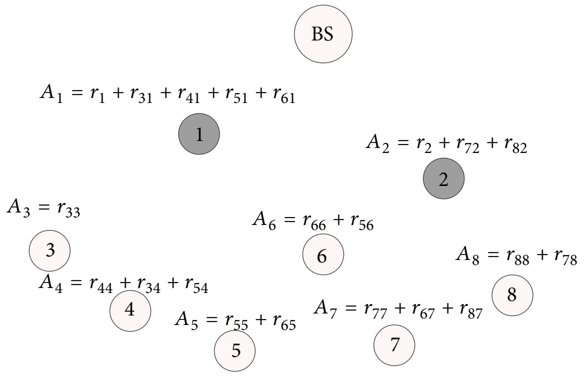

BS
$A_1 = r_1 + r_{31} + r_{41} + r_{51} + r_{61}$
1
$A_2 = r_2 + r_{72} + r_{82}$
2
$A_3 = r_{33}$
3
$A_6 = r_{66} + r_{56}$
6
$A_8 = r_{88} + r_{78}$
$A_4 = r_{44} + r_{34} + r_{54}$
8
4
$A_5 = r_{55} + r_{65}$
$A_7 = r_{77} + r_{67} + r_{87}$
7
5

Figure 4: Data mixing.

with its raw data piece r_{44}, and then the aggregation result of node 4 is $A_4 = r_{44} + r_{34} + r_{54}$.

When all the sent data pieces are received, leaf nodes begin to encrypt their aggregation results and send them to the aggregator nodes. The aggregators need to wait for a certain time to receive all the encrypted data pieces and then aggregate all the received data pieces with their own data. After the aggregation, the aggregators encrypt and send the aggregation results to their parent nodes. The process continues until all the aggregation results arrived at the base station. This procedure is described in Figure 5.

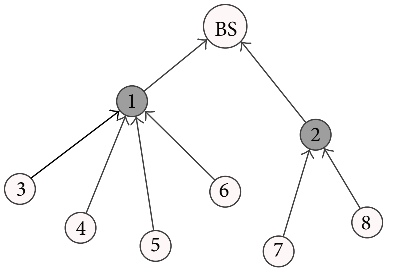


Figure 5: Data aggregation.

As we can see from the work process of ESMART, a leaf node needs to perform 2 encryption operations and 1 decryption operation during the process of data aggregation. For an aggregator node, only 1 encryption operation and 1 decryption operation are needed to be performed. Obviously, this degree of computational complexity has very little effect on the life cycle of sensor network and does not require a highly computational capacity. Meanwhile, compared to computational complexity, communication overhead has always been a major factor that affects the life cycle of sensor network. Thus, the analysis of the communication overhead of ESMART will be one of the focuses in Section 4.

4. Simulation and Analysis

In this section, simulations are carried out for analyzing and comparing the performance of TAG, SMART, and ESMART scheme. First, we set up the simulation environment in MATLAB which is an area of 400 m ∗ 400 m and there are 600 sensor nodes deployed in this area randomly with $P_a = 0.3$. J denotes the maximum number of data pieces that the primitive data can be sliced into in the network. Thus, the value of J will be set based on the demands of simulation. The other network parameters which are beyond the scope of this paper will not be discussed. The performance comparisons of the schemes are focused on three aspects: privacy preservation, communication overhead, and accuracy.

4.1. Privacy-Preservation Analysis of ESMART. In ESMART scheme, a sensor node which received a query from the base station will start a collection of target data and slice the collected data into j pieces and then encrypt $j-1$ pieces of the data and send them to the neighbor nodes. Therefore, the outdegree of the node is $j-1$. Meanwhile, the node will receive k pieces of encrypted data from its neighbors, so the indegree of the node is k. By analyzing the attack model we know that only in the case that all the $J-1$ outgoing and k incoming links are broken by the eavesdropper, the eavesdropper will be able to steal the private data of the node. In the SMART scheme, J is a fixed value which is different from ESMART scheme. Therefore, $P_s(q, J)$, the disclose probability of private data in SMART, can be expressed as below:

$$P_s(q, J) = q^{J-1} \sum_{k=0}^{\mathrm{id}_{\max}} P(\text{indegree} = k)\, q^k. \tag{2}$$

The variable q denotes the probability that a secure link of the node is broken. $\mathrm{id}_{\max}$ is the maximum indegree of nodes in the network and the value of $\mathrm{id}_{\max}$ depends on J. $P(\text{indegree} = k)$ is the probability that the indegree of the node equals k. $\sum_{k=0}^{\mathrm{id}_{\max}} P(\text{indegree} = k)q^k$ denotes the probability that all the outdegrees of the node are broken and q^{J-1} is the probability that all the indegrees of the node are broken. In the ESMART scheme, J is a preset value which denotes the maximum number of data pieces that the primitive data can be sliced into, and j is a variable whose range of values allowed is from 2 to J. Then the formula of the disclose probability of private data in ESMART can be described as below:

$$\begin{aligned} P_e(q, J) = {} & \sum_{j=2}^{J} P(\text{outdegree} = j)\, q^{j-1} \\ & \times \sum_{k=0}^{\mathrm{id}_{\max}} P(\text{indegree} = k)\, q^k, \end{aligned} \tag{3}$$

where $P(\text{outdegree} = j)$ is the probability that the outdegree of the node is j and then $\sum_{j=2}^{J} P(\text{outdegree} = j)q^{j-1}$ denotes the probability that all the outgoing links of the node are broken. As discussed in the previous chapter, the probability distribution of the node outdegree in the network depends on the performance-cost ratio of j which means the j with higher performance-cost ratio will have higher probability to

Table 1: The probability distribution of j with $J = 6$.

The value of j	2	3	4	5	6
probability	0.3448	0.2299	0.1724	0.1379	0.1150

be chosen as the outdegree of the node. The performance-cost ratio formula of j is as below:

$$\begin{aligned} f(j) &= \frac{1 - P_s(q, j)}{j} \\ &= \frac{1 - q^{j-1} \sum_{k=0}^{\mathrm{id}_{\max}} P(\text{indegree} = k)\, q^k}{j}, \end{aligned} \tag{4}$$

where $1 - P_s(q, j)$ is the secure probability of the communication link in the network and j denotes the communication cost of the node, according to the formula of communication overhead in ESMART. Therefore, the probability distribution formula of J with different values of j can be concluded as below:

$$F(J, j) = P(\text{outdegree}) = \frac{f(j)}{\sum_{j=2}^{J} f(j)}. \tag{5}$$

Table 1 illustrates the probability distribution of j when $J = 6$.

As we can see from Table 1, the larger value of j, the lower probability of the value to be selected. That is, the larger value of j, the lower performance-cost ratio of the scheme. After the mathematical derivation, the simulations are carried out to compare the privacy-preservation performance of SMART and ESMART. TAG scheme was excluded from the comparison due to lacking protection mechanism of data privacy. The two schemes are simulated in the same network environment with $J = 3$, $P_a = 0.3$, and the simulation results are shown in Figure 6.

As we can see from Figure 6, the disclose probability of the private data in ESMART scheme is lower than that in SMART. That is because the amount of communication in ESMART is smaller than that in SMART, which means less chance to eavesdrop on the data pieces of one node, that is, the disclose probability of the private data is reduced. Meanwhile, the introduction of variable j brings another advantage: the times of data slicing of a sensor node in the network are only known by itself, so even if a malicious node eavesdropped on several data pieces from the node, it is not sure that all the data pieces of the node are eavesdropped on, unless the number of eavesdropped data pieces equals J which is the theoretical maximum value of j. Thus, it is more difficult for the malicious nodes to steal private data. Given the above, the privacy-preservation performance of ESMART is better than that of SMART.

4.2. Communication Overhead Analysis. How to make a reasonable tradeoff between security and communication overhead is always a critical issue in the designing of data aggregation scheme for wireless sensor network. As we can see from the work in [4], SMART scheme introduced the technique of data slicing and mixing to achieve privacy preservation of wireless sensor network but also brings some additional communication overhead. In our proposed scheme ESMART, we decrease the communication overhead during the procedure of aggregation while keeping a good performance of privacy preservation. In this section, to validate the efficiency of ESMART scheme, we conclude the calculating formulas of the communication overhead of TAG, SMART, and ESMART and then deploy these aggregation schemes in the same network environment separately to avoid the influence of other factors on the simulation results. First, we assume a network consisting of N sensor nodes. In TAG scheme, the nodes upload their raw data directly without any protection of private data; thus the communication overhead can be expressed as below:

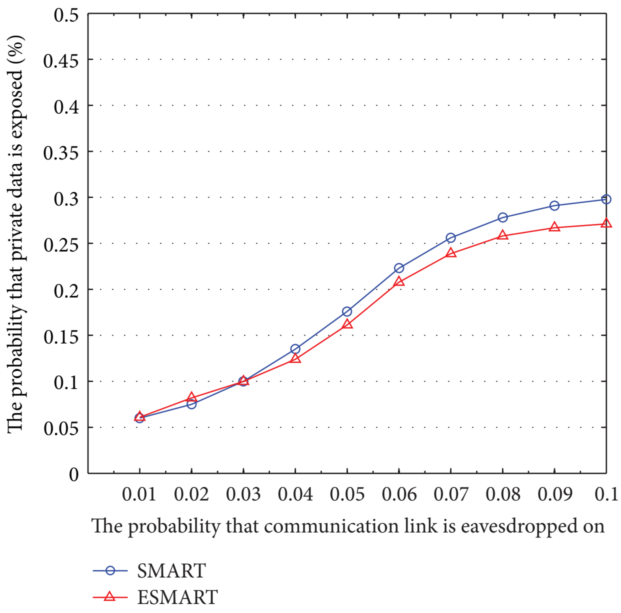


Figure 6: Privacy preservation of SMART and ESMART.

$$\text{TAG: } \mathrm{CO}_t = N. \tag{6}$$

In SMART scheme, each node slices the primitive data into J pieces and sends $J - 1$ pieces to its neighbor nodes to preserve privacy. After that, the nodes upload the processed data to their parent nodes according to the procedure of TAG scheme, so the communication overhead formula of SMART can be concluded as below:

$$\text{SMART: } \mathrm{CO}_s = N \cdot J. \tag{7}$$

In ESMART scheme, the aggregator nodes do not participate in data slicing. Thus, the data traffic of aggregator nodes in the network is $N \cdot P_a$, where P_a denotes the probability that a sensor node becomes an aggregator node. Therefore, there are $N(1-P_a)$ leaf nodes in the network. Each leaf node sends $j_i - 1$ pieces of raw data to the neighbor nodes and j_i denotes the value of j of node i. After that, it transmits the mixed data to the aggregator node. So, we can express the communication overhead formula of ESMART as below:

$$\text{ESMART: } \mathrm{CO}_e = N \cdot P_a + \sum_{i=1}^{N(1-P_a)} j_i, \quad (j_i \in [2, J]). \tag{8}$$

As we introduced in Section 3, j is a variable which takes values based on a probability distribution function $F(J, j)$ and the range of values allowed is from 2 to J. Therefore, we introduce function $F(J, j)$ into the calculating formula of CO_e and derived the formula which is shown below:

$$\mathrm{CO}_e = N \cdot P_a + \sum_{j=2}^{J} F(J, j) \cdot N(1 - P_a) \cdot j. \tag{9}$$

To compare the communication overheads of SMART and ESMART, we set $Q = \mathrm{CO}_s/\mathrm{CO}_e$. Thus, if $Q \geq 1$, then $\mathrm{CO}_s \geq \mathrm{CO}_e$. If $Q \leq 1$, then $\mathrm{CO}_s \leq \mathrm{CO}_e$. Therefore,

$$\begin{aligned} Q &= \frac{N \cdot J}{N \cdot P_a + \sum_{j=2}^{J} F(J, j) \cdot N(1 - P_a) \cdot j} \\ &= \frac{J}{P_a + (1 - P_a) \sum_{j=2}^{J} F(J, j) \cdot j}. \end{aligned} \tag{10}$$

From the definition of $F(J, j)$, we can deduce that $\sum_{j=2}^{J} F(J, j) \cdot j < J$, then we set $Q > Q'$, and the expression of Q' is shown as below:

$$\begin{aligned} Q > Q' &= \frac{J}{P_a + (1 - P_a) \cdot J} = \frac{J}{J + P_a - P_a \cdot J} \\ &= \frac{J}{J + P_a(1 - J)}. \end{aligned} \tag{11}$$

For $J \geq 2$ and $P_a \in (0, 1)$, we can determine that $P_a(1 - J) < 0$. Thus, $J + P_a(1 - J) < J$ and then $Q > Q' > 1$. From the derivation above, we can conclude that $\mathrm{CO}_s > \mathrm{CO}_e$. After the mathematical derivation, we deployed TAG, SMART, and ESMART scheme in the simulation environment separately and the value of J in the simulation environment is set to 4, $P_a = 0.3$. The simulation results are shown in Figure 7.

It is concluded from the simulation results above that because of lacking protection of data privacy, the number of messages transmitted during the procedure of data acquisition and aggregation under the TAG scheme is very small. Meanwhile, owing to the introduction of data slicing technique, the communication overhead of the network deployed with SMART is much higher than that with TAG. And the communication overhead of the proposed scheme ESMART is about 37% to 46% lower than that of SMART which means that ESMART is more energy efficient than SMART while keeping a better performance of privacy preservation.

4.3. Accuracy Analysis. The accuracy of aggregation result is an important indicator of the performance of data aggregation scheme. Theoretically, the accuracy of aggregation result is 100% which means the final aggregation result equals the sum of the collected data of every node in the network. In reality, due to channel sharing character of wireless network, data loss is inevitable during the procedure of data aggregation. The more transmissions per unit time in the network, the more collisions and then the more data loss. In other words, the data accuracy in wireless network is inversely associated with the communication overhead. Thus, from the simulation results of communication overhead of the schemes, we can theoretically deduce that the accuracy performance of TAG is the best in these schemes, secondly ESMART, thirdly SMART. After the analysis, we deployed the schemes in the simulation environment with $J = 3$, $P_a = 0.3$, and the simulation results are shown in Figure 8.

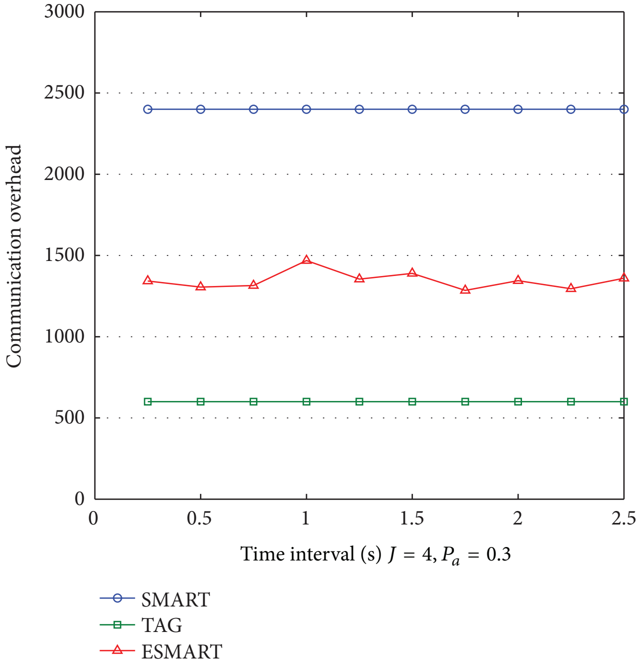


Figure 7: Communication overhead of TAG, SMART, and ESMART.

As we can see from the simulation results, due to the low amount of communication, the accuracy of TAG scheme is about 90% when the interval time is long enough. Instead, the accuracy of SMART scheme is approximately 67% which is caused by the large amount of data traffic. The proposed scheme ESMART can provide a better accuracy of 83% than SMART while keeping a good performance of privacy preservation. The simulation results confirmed the previous deduction that the aggregation scheme with low communication overhead can achieve a good performance of data accuracy.

5. Conclusions

Wireless sensor network consists of a large number of low-powered, resource-constrained sensor nodes and is usually deployed in unattended environment. Sensor nodes in the network can sense specific information and perform wireless communication within a small range of their location. To

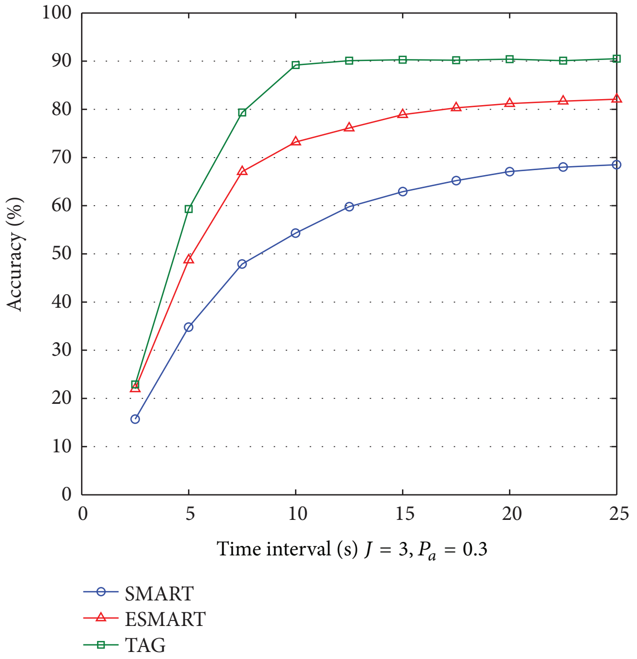


Figure 8: Accuracy of TAG, SMART, and ESMART.

reduce the energy consumption of the nodes during the transmission, data aggregation scheme has been widely used in WSN. Meanwhile, how to provide the protection of data privacy during the aggregation is becoming a desiderative problem to be solved.

In this paper, we proposed an energy-efficient secure data aggregation scheme ESMART. The simulation results and analysis proved that whether in energy saving ability or data accuracy ESMART scheme is better than SMART scheme while keeping a good performance of privacy preservation which is not provided in TAG scheme. It is an efficient secure data aggregation scheme with a value of practical application and it achieved the expected design requirements. In future research work, we will focus on the designing of secure data aggregation scheme with the ability to guarantee data integrity.

Acknowledgments

This research is supported by Beijing Science and Technology Program under Grant Z121100007612003, Beijing Natural Science Foundation under Grant 4132057, and National Natural Science Foundation under Grant 61371071.

References

[1] S. Madden, M. J. Franklin, J. M. Hellerstein, and H. Wei, "TAG: a Tiny AGgregation service for ad-hoc sensor networks," in *Usenix Association Proceedings of the 5th Symposium on Operating Systems Design and Implementation*, pp. 131–146, 2002.

[2] 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.

[3] K. J. Mukesh and T. Sharma, "Secure data aggregation in wireless sensor network: a survey," *International Journal of Engineering Science and Technology*, vol. 3, no. 3, pp. 2013–2019, 2011.

[4] W. He, X. Liu, H. Viet, K. Nahrstedt, and T. Abdelzaher, "PDA: privacy-preserving data aggregation for information collection," *ACM Transactions on Sensor Networks*, vol. 8, no. 1, article no. 6, 2011.

[5] W. He, H. Nguyen, 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)*, vol. 1–7, pp. 4071–4077, November 2008.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] A. Miyaji and K. Omote, "Efficient and secure aggregation of sensor data against multiple corrupted nodes," *IEICE Transactions on Information and Systems*, vol. E94-D, no. 10, pp. 1955–1965, 2011.

[11] H. Alzaid, E. Foo, J. G. Nieto, and E. Ahmed, "Mitigating On-Off attacks in reputation-based secure data aggregation for wireless sensor networks," *Security and Communication Networks*, vol. 5, no. 2, pp. 125–144, 2012.

[12] S. ZareAfifi, R. Verma, B. King, P. Salama, and D. Kim, "Secure countermeasures to data aggregation attacks on sensor networks," in *Proceedings of the 55th IEEE International Midwest Symposium on Circuits and Systems (Mwscas '12)*, pp. 856–859, 2012.

[13] W. T. Zhu, F. Gao, and Y. Xiang, "A secure and efficient data aggregation scheme for wireless sensor networks," *Concurrency Computation Practice & Experience*, vol. 23, no. 12, pp. 1414–1430, 2011.

[14] X. M. Dong and S. S. Li, "Secure data aggregation approach based on monitoring in wireless sensor networks," *China Communications*, vol. 9, no. 6, pp. 14–27, 2012.

[15] F.-Y. Lei, C. Fu, X. Li, and J. Chen, "Secure data aggregation solution based on dynamic multiple cluster key management model," *Journal of Internet Technology*, vol. 12, no. 3, pp. 465–476, 2011.

[16] H. M. Sun, C. H. Chen, and P. C. Li, "A lightweight secure data aggregation protocol for wireless sensor networks," *International Journal of Innovative Computing Information and Control*, vol. 8, no. 10A, pp. 6503–6514, 2012.

[17] A. S. Poornima and B. B. Amberker, "Secure end-to-end data aggregation (SEEDA) protocols for wireless sensor networks," *Ad Hoc & Sensor Wireless Networks*, vol. 17, no. 3-4, pp. 193–219, 2013.

[18] 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.

[19] Q. Xie, "A new authenticated key agreement for session initiation protocol," *International Journal of Communication Systems*, vol. 25, no. 1, pp. 47–54, 2012.

[20] D.-C. Lou and H.-F. Huang, "Efficient three-party password-based key exchange scheme," *International Journal of Communication Systems*, vol. 24, no. 4, pp. 504–512, 2011.

[21] D. He, J. Chen, and J. Hu, "A pairing-free certificateless authenticated key agreement protocol," *International Journal of Communication Systems*, vol. 25, no. 2, pp. 221–230, 2012.

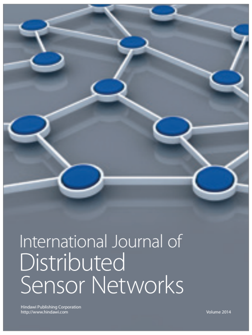
The Scientific World Journal

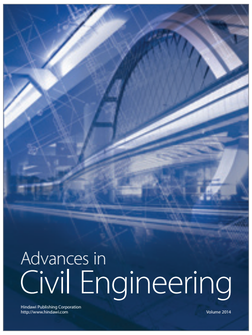
International Journal of
Rotating Machinery

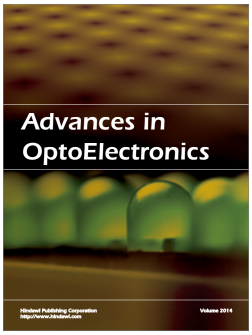
Journal of
Engineering

Journal of
Sensors

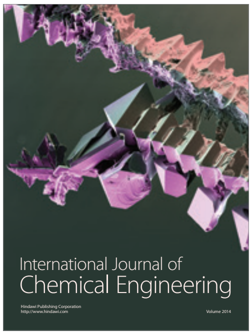
Advances in
Mechanical Engineering


International Journal of
Distributed Sensor Networks


Advances in
Civil Engineering

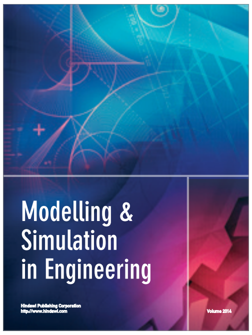

Advances in
OptoElectronics

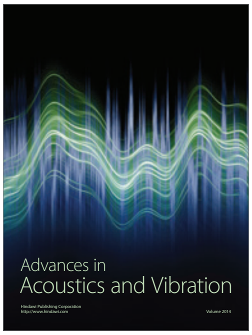
Journal of
Robotics


International Journal of
Chemical Engineering

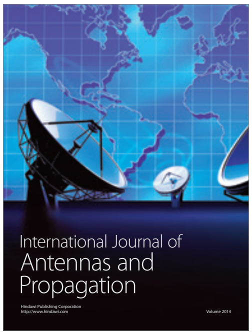
VLSI Design

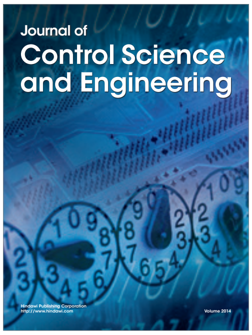
International Journal of
Navigation and Observation

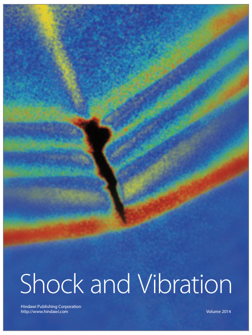

Modelling &
Simulation
in Engineering


Advances in
Acoustics and Vibration

Active and Passive
Electronic Components


International Journal of
Antennas and Propagation


Journal of
Control Science and Engineering


Shock and Vibration

Journal of
Electrical and Computer Engineering