Research Article

Data Fusion Algorithm of Privacy Protection Based on Qos and Multilayers Hierarchically

Li Li, Qin Qin, Li Hua, and Li Jian

Department of Computer Science and Engineering, Henan Institute of Engineering, Zhengzhou 451191, China

Correspondence should be addressed to Li Li; ll_mm2005@163.com

Received 21 September 2013; Accepted 19 November 2013

Academic Editor: Ying-Hong Wang

Copyright © 2013 Li 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.

Because of the special nature of wireless sensor networks, data fusion process is vulnerable to attacks by the destroyer, which is useful for reducing the network communication overhead and improving the data transmission efficiency. Therefore, it is very essential to provide transmission with high security and low energy consumption for sensitive data fusion in wireless sensor network. In this paper, a data fusion algorithm of privacy protection based on Query of Server (Qos) and hierarchical multilayers was put forward, which divided the required privacy protection levels according to different safety requirements, and set up hierarchical network models. In data fusion process, delay constraint was added to guarantee service quality of Qos, and this would reduce energy consumption overhead. Meanwhile, it would guarantee the accuracy of the data and reduce the probability of the whole network information exposed.

1. Introduction

Wireless sensor network (WSN) data fusion technique has broad application prospects, which is one of the key technologies in the internet of things. Data fusion technology is the process of processing multiple copies of data and it gets more effective data to users. But wireless sensor network is vulnerable to various attacks for its open and self-organizing features. The fusion data attacks may cause bad results, which not only makes the wireless sensor network lose the original construction objective but also may cause much more damage. Therefore, it is very necessary to provide transmission of security and low energy consumption for sensitive data fusion in wireless sensor network.

Currently, data fusion safe programs can be divided into two kinds. One is the program based on data integrity, and the other is data confidentiality [1]. In traditional ways, Hop-by-hop encryption is the main way to protect the confidentiality [2]. The confidentiality of end-to-end and authentication of data fusion technology are the two main goals of safety data fusion protocol [3]. Data security in sensor networks fusion scheme was proposed in the paper in [4], and secure data fusion in wireless sensor network was proposed based on the homomorphic MAC in [5] by Wei et al. The security of data fusion was studied by both of them in wireless sensor network [4, 5].

However, security will increase the resource cost and reduce the efficiency improved by data fusion mechanism. Therefore, when the effective security of data fusion, calculation, and transmission was set up, resource consumption should be reduced as little as possible, and network life cycle should be prolonged as long as possible [6]. In this paper, a data fusion model of privacy protection based on Qos and hierarchical multilayers was put forward. The core problem of privacy protection in data fusion process is how to achieve data fusion of the privacy accurate and efficient ratio. To solve this problem, researchers have proposed some privacy protection algorithms [7–10]. The earliest is the PDA algorithm by He et al. to data fusion of privacy [11], and it includes two kinds of privacy protection data fusion algorithms. One is fusion algorithm of CPDA of privacy protection data based on clustering, and the other is fusion algorithm of SMART of privacy protection data based on distributed way.

Based on the problems above, a data fusion algorithm of privacy protection based on Qos and hierarchical multilayers was put forward, which divided the required privacy protection levels according to different safety requirements and set
up hierarchical network models. In data fusion process, delay
constraint was added to guarantee service quality of QoS, and
this would reduce energy consumption overhead. Meanwhile,
its would guarantee the accuracy of the data and reduce the
probability of the whole network information exposed. The
main structure of this paper is as follows: the first part gives
the current situation at home and abroad; the second part
pots forward the system structure of data fusion in wireless
sensor networks; the third part puts forward the multilayers
hierarchically data fusion approach to privacy protection; the
fourth part gives the using method of the QoS applied in
mulilayers hierarchically data to the privacy protection; the
fifth part is the experiment of a modern vegetable planting
base to validate the effectiveness of the method.

2. System Structure of WSN in Data Fusion

Data fusion is a subset of information fusion, which can be
used for processing multiple copies of data or information.
More effective and useful data to user can be combined [1].
For WSN, the data fusion technology can greatly reduce
the amount of data transmission to WSN and reduce the
data conflict. It can reduce the network congestion and save
costs effectively, too, which will prolong the lifetime of
network.

Sensor nodes are divided into ordinary nodes, fusion
nodes, and sink nodes, by WSN in data fusion process. Effect
of data fusion technique was studied through two kinds of
methods with theoretical analysis and simulation test. The
results show that the minimum energy consumption ratio
of using data fusion technology and not using data fusion
technology is as follows:

\[
\lim_{d \to \infty} \frac{N_D}{N_A} = \frac{1}{k}.
\]  \hspace{1cm} (1)

In this equation, \(N_D\) is the data transmission frequency
of data fusion and \(N_A\) is the data transmission frequency
of no data fusion. Meanwhile, \(d\) is distance between the sensor
nodes and the sink node and \(k\) is the quantity of the data
source node. This equation shows that the more data source
network node has, the more energy will be saved in data
fusion process. Meanwhile, if the size of the network is larger,
the energy saving of data fusion is much more significant.
This shows that the data fusion plays an important role in
saving energy of sensor node.

3. Data Fusion of Privacy Protection Method
   Based on Multilayers Hierarchically

3.1. Layered Data Fusion. Wireless sensor network consists
of a large number of sensor nodes, which are deployed in
the monitoring region. Sensor nodes are divided into three
categories. It includes base station, cluster head, and
the common sensor nodes. The base station has enough
energy and rich resources, and cluster head has less of them.
However, energy and resources of ordinary sensor node are the
least [11].

In the model of wireless sensor network, a node can
establish a communication link with any other node in
same clusters through different sharing keys. Each cluster of
wireless sensor network has \(n\) nodes, one of which is a cluster
head, and the others are ordinary sensor nodes [12]. This
paper proposes a sequence flow diagram of the calculation
procedure of data fusion, which is shown in Figure 1.

3.2. Data Fusion of Classification Privacy Protection. In data
fusion of classification privacy protection, in order to reduce
the communication, according to the barrel principle of
information security, privacy level packet contains a mini-
num of three node groups. Secondly, next packet contains
four nodes, and so on. If all nodes in cluster are divided into a
group, this situation is defined as the highest level of privacy.

For a cluster containing sensor node number of \(m\), if it
can be divided into two privacy levels, its highest level of
privacy is the same as the aggregation scheme of private data
based on cluster. At the lowest level of privacy, \(m\) sensor nodes
are divided into two pretreatment groups. One contains three
sensor nodes, and they are \(A, B,\) and \(C\). Their private data are
respectively \(a, b,\) and \(c\). The other contains four nodes, which
are \(E, F, G,\) and \(H\), and their private data are \(e, f, g,\) and \(h,\)
respectively.

For the three members of the group, the pretreatment
process is similar to the private data based on cluster
aggregation scheme. Pretreatment values can be obtained by
nodes \(A, B,\) and \(C\) as follows:

\[
F_A = v_A^A + v_A^B + v_A^C = (a + b + c)
+ r_{11}x_A + r_{12}x_A^2,
\]
\[
F_B = v_B^A + v_B^B + v_B^C = (a + b + c)
+ r_{11}x_B + r_{12}x_B^2,
\]
\[
F_C = v_C^A + v_C^B + v_C^C = (a + b + c)
+ r_{11}x_C + r_{12}x_C^2.
\]

In the above expression, \(r_{11} = r_1^A + r_1^B + r_1^C, r_{12} = r_2^A + r_2^B + r_2^C.

For the four members of the group, \(E, F, G,\) and \(H,\)
respectively, we use common nonzero digits \(x_E, x_F, x_G,\) and
\(x_H\) as seeds, and then node \(E\) will be calculated as follows:

\[
v_E^E = e + r_1^EX_E + r_2^EX_E + r_3^EX_E^2,
\]
\[
v_E^F = e + r_1^EX_F + r_2^EX_F + r_3^EX_F^2,
\]
\[
v_E^G = e + r_1^EX_G + r_2^EX_G + r_3^EX_G^2,
\]
\[
v_E^H = e + r_1^EX_H + r_2^EX_H + r_3^EX_H^2.
\]  \hspace{1cm} (3)

In the above expression, \(r_1^E, r_2^E,\) and \(r_3^E\) are random
numbers generated by node \(E.\) Similarly, nodes \(F, G,\) and \(H\)
are calculated to get \(v_F^E, v_F^F, v_F^G, v_F^H, v_G^E, v_G^F, v_G^G, v_G^H, v_H^E, v_H^F, v_H^G,\)
and \(v_H^H.\) The node \(E\) uses secret key shared with the nodes
of $F$, $G$, and $H$, respectively to encrypt $v^E_F$, $v^E_G$, and $v^E_H$, and send the results to $F$, $G$, and $H$, respectively. In the same way, the node $F$ encrypts $v^F_E$, $v^F_G$, and $v^F_H$ and sends the results to $E$, $G$, and $H$, respectively. The node $G$ encrypts $v^G_E$, $v^G_F$, and $v^G_H$ and sends the results to $E$, $F$, and $G$, respectively.

After receiving $v^E_F$, $v^E_G$, and $v^E_H$, node $E$ starts calculating pretreatment values as follows:

$$F_E = v^E_F + v^E_G + v^E_H = (e + f + g + h) + r_{21}x_E + r_{22}x_G^2 + r_{23}x_H^3. \tag{4}$$

In the above expression,

$$r_{21} = r_1 + r_1^* + r_1^F + r_1^H,$$

$$r_{22} = r_2 + r_2^* + r_2^F + r_2^H,$$

$$r_{23} = r_3 + r_3^* + r_3^F + r_3^H. \tag{5}$$

Pretreatment values of nodes $F$, $G$, and $H$ are calculated in the same way.

Then nodes $A$, $B$, $C$, $D$, $E$, $F$, $G$, and $H$ send $F_A$, $F_B$, $F_C$, $F_D$, $F_E$, $F_F$, $F_G$, and $F_H$, respectively. After the cluster head node receives $F_A$, $F_B$, $F_C$, $F_D$, $F_E$, $F_F$, $F_G$, and $F_H$, data fusion process is performed and equation will be constructed as

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}^{-1} \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}. \tag{6}$$

In the above expression,

$$U_1 = \begin{bmatrix} a + b + c \\ r_{11} \\ r_{12} \end{bmatrix}, \quad U_2 = \begin{bmatrix} e + f + g + h \\ r_{21} \\ r_{22} \\ r_{23} \end{bmatrix},$$

$$G_1 = \begin{bmatrix} 1 & x_A^2 \\ 1 & x_B^2 \\ 1 & x_C^2 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 1 & x_E^3 \\ 1 & x_F^3 \\ 1 & x_G^3 \end{bmatrix},$$

$$F_1 = [F_A \ F_B \ F_C]^T,$$

$$F_2 = [F_E \ F_F \ F_G \ F_H]^T.$$

The first element of $U_1$ can be added to the first element of $U_2$ to get the data fusion value.

4. Application of Qos in Data Fusion of Privacy Protection Based on Multilayers

4.1. The Service Quality of Qos in WSN. Three basic design goals of mechanism research of Qos in wireless sensor network are real time, reliability, and validity of resource utilization. Metrics definition, transmission mode of non-end-to-end data, and node resources are the main study problems of Qos wireless sensor networks [13].

4.2. Multitiered Privacy Protection Data Fusion Protocol Based on Qos. In WSN, Multitiered privacy protection data fusion protocol based on Qos (MPPDF-Qos) includes three kinds of nodes, which are Qos node, fusion node, and leaf node. In the traditional technology of data fusion, Qos node is the root of
the data fusion structure tree to get the final result of data fusion. 
The merging data of fusion node receives from its child nodes 
and collected data by itself and sends to the parent node. Leaf 
node is responsible for collecting data and sending it to the 
parent node.

In this algorithm, we assume that Qos is a Time Delay 
(TiD), and node is assigned a Communication Delay (ComD) 
and a Fusion Delay (FusD). All the nodes communicate in 
ComD time, and do transmission and fusion process in the 
FusD time period:

\[
\text{ComD} = \frac{\text{TiD} \cdot \text{MinDeg}}{\text{MinDeg} + \text{MaxDEPTH}}, \\
\text{FusD} = \frac{\text{TiD} \cdot \text{MaxDEPTH}}{\text{MinDeg} + \text{MaxDEPTH}}.
\] (8)

In the above expression, MaxDEPTH represents the 
maximum depth of the tree data fusion in wireless sensor 
networks and MinDeg is the minimum required entry node 
degrees.

4.2.1. The Algorithm Steps

The First Step: preparatory work. Each node \(i \ (i = 1, \ldots, N)\) 
selects a set of nodes \(S_i \ (|S_i| = \text{MinDeg})\). In time of TiD, TAG 
algorithm is used to get data fusion tree. The number of child 
nodes is recorded, in the data fusion process at each node of 
\(i \ (i = 1, \ldots, N)\), which is denoted as deg. If there are leaf 
node obviously, \(\text{deg}_i = 0\). At this time, one result is obtained, 
and the cycle number was increased by one by the node.

The Second Step: nodes collusion. The size of MinDeg is 
used to determine how much the allocation of time slice, 
and \(t_s(j = 1, \ldots, \text{MinDeg})\) is recorded as time slice in 
ComD. The values are distributed uniformly. Therefore, \(t_p = 
(\text{ComD}/\text{MinDeg}), \ (p = 1, \ldots, \text{MinDeg})\).

Firstly, the network node is assigned within the ComD 
and time slice of \(t_1\), and the nodes deg, with the values 
are one will collude with other nodes. It means that leaf deg 
ode node chooses time randomly in the allocated time slice of \(t_1\). Then, 
in the allocated time slice of \(t_2\), the nodes deg, with the values 
are one will collude with other nodes until node with the deg 
of MinDeg − 1 colludes with other nodes in the allocated time 
slice of MinDeg.

4.2.2. Collusion Method

(1) Nodes firstly verify their values of deg, and 
determine its size relationship with MinDeg. If the 
value is smaller than MinDeg, they only need to 
receive data within the ComD, and it is not necessary 
to send collusion data. After ComD, it goes directly to 
third step. If the value is larger than MinDeg, node 
\(j\) is chosen from node \(i\), and \(\text{Seed}_{ij}\) is sent to node 
\(j\). Meanwhile, one is added to deg, and \(\text{Seed}_{ij}\) is 
subtracted from the node data. Seeds are produced 
by node \(i\) using a random number, range of data, and 
calculation of MinDeg:

\[
\text{Seed}_{ij} = \frac{r \cdot (V_{\text{superbound}} - V_{\text{lowerbound}})}{\text{MinDeg}}. \quad (9)
\]

(2) When the node \(i\) receives the encryption seeds, it 
uses the sharing key to decrypt with neighbor nodes 
and obtain \(\text{Seed}_{ij}\). Then, the obtained seeds are added 
with data recorded by themselves, and deg, is added 
by one. The pseudodata obtained is \(f_i + \text{Seed}_{ij}\). Then 
it starts step (1) jumped from step (2).

The Third Step: fusion process. In the FusD time, fusion steps 
of TAG algorithm are used to do fusion process from the 
bottom node to the top node based on the data fusion tree 
set up in the first step, and data fusion result will be obtained 
at Qos at last. At this time, the node cycle number will be plus 
one. If loop is smaller than Loop, step two will be executed. 
Otherwise, the algorithm will end.

5. Verification of Case

Wireless sensor networks have been widely applied in the 
field of intelligent agriculture, and experiments were con-
ducted in a modern vegetable planting base in city. Whether 
the irrigation of vegetables was needed is mainly determined 
by the detection of air humidity. For the comprehensive 
analysis of the data fusion scheme, it was necessary to analyze 
the data accuracy, security, and data transmission overhead.

5.1. The Experimental Platform. According to the character-
istics of data fusion in wireless sensor network, Matlab sim-
ulation platform was used to analyze the data. In the region 
of 50 m × 100 m for vegetable greenhouse, thirty sensor nodes 
were deployed, and the transmission radius of each node was 
15 m. Ten vegetable bases like this were selected and tested.

5.2. System Model. Network experiment model was as shown 
in Figure 2, and humidity sensor was AM1001, for sensing 
the humidity in the greenhouse. The fusion node first did 
security data fusion to the received information, and then 
it transmitted the information to the base station. The 
traditional methods of data fusion and MPPDF-Qos data 
fusion methods were used to do data fusion of perceived 
humidity data and the final results of the analysis were as a 
basis whether to irrigate vegetable.

5.3. Experimental Results and Analysis. In the MPPDF-Qos 
scheme, each node could establish links with other nodes 
through different shared keys. Probability on violation of 
privacy by eavesdropping was zero; therefore, only the coor-
dinated attack situation should be considered. In order to 
prevent the attack, each group needed to have at least three 
sensor nodes. Therefore, the definition of privacy level for 
minimum packet was the packet with three sensor nodes. So, 
we set \(J \geq 3\) when calculating the allocation of time slice \(t_p\).
and in order to reflect the actual value of the data fusion, the delay time TiD must be set to more than 30 seconds. The value of the random number r is evenly distributed in (−1, +1).

1. **The Energy Consumption.** In the MPPDF-Qos scheme, nodes communicated only within the same group, in addition to transmitting preprocessing data to the cluster head. However, in the other scheme, all nodes in the cluster could have communication. Therefore, the MPPDF-Qos scheme saved data overhead. Figure 3 was comparison of data traffic of the CPDA scheme, SMART scheme, and MPPDF-Qos scheme.

2. **Security.** In Figure 4, privacy protection performances of the MPPDF-Qos scheme and CPDA scheme were compared when there were 30 nodes in a cluster. From the figure, it can seem that it would increase with the increasing rate of node capture in three schemes, and the probability of exposure of data information had been improved. However, it could provide privacy protection based on different privacy levels required in MPPDF-Qos scheme. Even if it was intercepted for the same data as CPDA, the probability is that the whole information exposed was much lower than in the CPDA.

3. **Accuracy.** In order to reduce the probability of data privacy exposed, a random number would be generated in the course calculating of cluster by nodes. Disturbance data would be added in the transmitted data. Data perturbation could reduce the probability of data exposed, but it often affected the accuracy of the results.

SMART needed to send more count data to its neighbor node; thus, it would cause more collision to reduce the accuracy. Data collision chance was avoided in MPPDF-Qos scheme, so its data accuracy was ideal. There are 10 simulations in each numerical value for each EpochD and the means of these simulation results are as our result. Accuracy analysis was shown in Figure 5.

### 6. Discussion and Conclusion

A data fusion algorithm of privacy protection based on Qos and multilayers hierarchically was put forward, which
divided the required privacy protection levels according to the different safety requirements, and hierarchical network models were set up, too. In data fusion process, delay constraint was added to guarantee service quality of Qos. This scheme has great advantages in energy consumption and safety. In the next period of time, data accuracy will be the key point to study.

Wireless sensor network data fusion technology is an important part of the internet of things, and it has wide application prospect in real life. How to provide security transmission to sensitive data fusion in wireless sensor network and how to protect its privacy are important technical support issues for the practical application.

However, security is often based on the overhead of the additional overhead, and too much overhead will reduce the efficiency of data fusion mechanism. Therefore, building an effective guaranteed security of data fusion, calculation, and transmission is important. Meanwhile, reducing resource consumption and prolonging the network life are necessary, too. It can meet the accuracy and privacy requirements of data fusion through MPPDF-Qos scheme and has lower energy consumption, too.

Conflict of Interests

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

Acknowledgments

This work has been supported by the National Natural Science Foundation of China (no. 61301232) and the Science Technologies Foundation of Henan, China (no. 13A520148).

References

[1] K. Jian, Z. Zongxian, T. Liwei et al., "Wireless sensor network data fusion technology," Computer Science, vol. 37, no. 4, pp. 31–35, 2010.
[2] Y. Yang, X. Wang, S. Zhu, and G. Cao, "SDAP: a secure hop-by-hop data aggregation protocol for sensor networks," ACM Transactions on Information and System Security, vol. 11, no. 4, article 18, 2008.
[3] C. Castelluccia, A. C.-F. Chan, E. Mykletun, and G. Tsudik, "Efficient and provably secure aggregation of encrypted data in wireless sensor networks," ACM Transactions on Sensor Networks, vol. 5, no. 3, pp. 1–36, 2009.
[4] Z. Peng, Y. Jianping, and L. Hongwei, "Data security in sensor networks fusion," Computer Science, vol. 38, no. 8, pp. 106–108, 2011.
[5] Q. Wei, S. Zhang, X. Hu, and X. Qin, "Secure data aggregation scheme based on homomorphic MAC for wireless sensor networks," Chinese Journal of Sensors and Actuators, vol. 24, no. 12, pp. 1750–1755, 2011.
[6] Q. Chen, Z. Tang, Y. Li, Y. Niu, and J. Mo, "Research on encryption algorithm of data security for wireless sensor network," Journal of Computational Information Systems, vol. 7, no. 2, pp. 369–376, 2011.
[7] S. Fan, D. Li, J. Zhao, and W. An, "Mine multi-sensor maximum likelihood estimation data fusion algorithm," Journal of Information & Computational Science, vol. 10, no. 12, pp. 3809–3814, 2013.
[8] M. Conti, L. Zhang, S. Roy, R. Di Pietro, S. Jajodia, and L. V. Mancini, "Privacy-preserving robust data aggregation in wireless sensor networks," Security and Communication Networks, vol. 2, no. 2, pp. 195–213, 2009.
[9] R. Bista, K.-J. Jo, and J.-W. Chang, "A new approach to secure aggregation of private data in wireless sensor networks," in Proceedings of the 8th IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC ’09), pp. 394–399, December 2009.
[10] R. Bista, K. Hee-Dae, and J.-W. Chang, "A new private data aggregation scheme for wireless sensor networks," in Proceedings of the 10th IEEE International Conference on Computer and Information Technology, pp. 273–280, July 2010.
[11] L. Li and K. Mu, "A data fusion method of wireless sensor network based on security protection," Journal of Computational Information Systems, vol. 9, no. 17, pp. 7029–7036, 2013.
[12] W. 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, May 2007.
[13] S. A. Munir, W. B. Yu, B. Ren, and J. Ma, "Fuzzy logic based congestion estimation for QoS in wireless sensor network," in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC ’07), pp. 4339–4344, March 2007.