doi:10.3969/j.issn.1000-436x.2014.z2.015

低能耗的隐私数据安全融合方法

谷勇浩[1]，郭达[2]，林九川[3]
（1. 北京邮电大学 计算机学院，北京 100876；
2. 北京邮电大学 电子工程学院，北京 100876；3. 公安部第三研究所，上海 201204）

摘 要：为解决物联网安全数据融合过程中，数据隐私保护与节点计算能力及能量受限之间的矛盾，在对现有方法优缺点分析的基础上，提出一种低能耗的隐私数据安全融合方法(LCSDA, low energy-consuming secure data aggregation)，该方法根据最短路径原则选择邻居节点，并且采用 Prim 最小生成树算法建立簇内数据融合路径。仿真结果表明，该方法可以有效降低节点能耗和簇头节点被捕获的概率，同时保证节点数据的隐私性。
关键词：物联网；隐私保护；安全数据融合；最小生成树算法；低能耗
中国分类号：TP393 **文献标识码**：A **文章编号**：1000-436X(2014)Z2-0112-05

Energy-saving privacy data secure aggregation method

GU Yong-hao[1], GUO Da[2], LIN Jiu-chuan[3]
(1. School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China;
3. The Third Research Institute of Ministry of Public Security, Shanghai 201204, China)

Abstract: For the Internet of things(IoT) secure data aggregation issues, data privacy-preserving and limited computation ability and energy of nodes should be tradeoff. Based on analyzing the pros-and-cons of current works, a low energy-consuming secure data aggregation method (LCSDA) was proposed. This method uses shortest path principle to choose neighbor nodes and generates the data aggregation paths in the cluster based on prim minimum spanning tree algorithm. Simulation results show that this method could effectively cut down energy consumption and reduce the probability of cluster head node being captured, in the same time preserving data privacy.
Key words: Internet of things; secure data aggregation; privacy preserving; minimum spanning tree algorithm; energy saving

1 引言

物联网已经在很多领域得到广泛应用，无线传感器网络是物联网的重要组成部分，它的功能是依靠分散在环境中的大量节点, 收集有用信息，以便人们使用这些信息进行分析与处理。数据融合技术是指对按时序获得的若干采集信息，在一定准则下加以自动分析、综合，以完成所需的决策和评估任务而进行的信息处理技术[1]。

数据融合过程中存在隐私泄露和其他安全问题，传统的安全数据融合方法多采用密码算法或安全路由协议来实现，但是传感器节点有限的计算能力、通信能力和存储空间限制了很多方法的使用[2]。因此，有效的安全数据融合方法，需要均衡数据的隐私保护需求与节点计算能力及能耗受限之间的矛盾[3]。

本文提出一种低能耗的隐私数据安全融合协议 LCSDA，该协议主要针对 CPDA[4]和 LCCPDA[5]

收稿日期：2014-11-04
基金项目：国家自然科学基金资助项目（61173017）；工信部通信软科学基金资助项目（2014-R-42）；信息网络安全公安部重点实验室开放课题基金资助项目（C14613）
Foundation Items: The National Natural Science Foundation of China (61173017); Communication Soft Science Foundation of Ministry of Industry and Information (2014-R-42); Key Lab of Information Network Security Foundation of Ministry of Public Security (C14613)

协议进行了改进。LCSDA 通过最小生成树算法构建簇内融合树，可以降低节点能耗，降低簇头节点被捕获的概率，同时保证了节点数据的隐私性。

2　相关工作

目前，安全数据融合方法主要包括如下内容。

1) 基于模式码的融合方法

文献[6]针对无线传感器网络提出一种基于模式码的低能耗安全数据融合算法 ESPDA，该算法在数据聚合过程中使用模式码标识并对节点原始数据进行分类。该算法具有降低节点能耗的优点，缺点是不适用于大规模网络，而且融合结果不精确。

文献[7]在模式码融合方法的基础上提出基于参照值的安全数据融合算法 SRDA。该算法不将节点原始数据直接发送到汇聚节点，而将采集的原始数据与初始设定的参考数据（即节点融合数据的平均值）进行差值运算，然后将该差值发送到汇聚节点。该方法降低数据通信量的同时提高了数据融合效率。但是，网络的中间节点不执行数据融合操作，使得节点能耗无法进一步降低。

2) 基于同态加密的融合方法

文献[8]中提出隐藏的数据融合算法 CDA，适用于各种物联网及传感器网络。该算法首先将各传感器节点的数据分成 n 份，并分别乘以密钥后，将密文发送到汇聚节点。汇聚节点对密文做模加运算后将结果发送给 Sink 节点，Sink 节点解密得到数据融合结果。CDA 算法中节点计算量小，但是数据的分片导致传输开销增大。

文献[9]提出 CMT 算法，该算法假定每个节点都与汇聚节点共享一个密钥。节点将原始数据与密钥做模加运算，对加密后的密文进行加法数据融合。相比 CDA 算法，CMT 算法传输开销小，但安全性有所下降。

3) 基于数据分割技术的融合方法

文献[4]提出基于簇的隐私数据融合算法 CPDA，虽然该方法提供较高的隐私保护性能，但是节点计算量大且簇内节点通信能耗高。这些不足正是本文需要解决的问题。文献[4]还提出一种基于数据分片的隐私保护数据融合算法 SMART。该方法将采集的数据分片，分片数据加密后按照不同的路径传输。因此，对于隐私攻击者来说，只有获得所有数据分片才能得到最终的数据隐私信息，这是很困难的事。但是，该算法的通信开销很大，节点能耗高。

针对 SMART 算法的上述不足，EEHA 算法[10]和 ESPART 算法[11]从减少数据通信量和提高精确度方面做了改进。EEHA 算法只让融合树中的叶子节点在分片后上传数据，因此网络数据通信量下降的同时降低了节点能耗。ESPART 算法给每个节点分配随机时间片，降低节点间碰撞的概率，同时减少节点间的数据传输量。

4) 基于完整性的融合方法

该类中有代表性的方法包括：安全可靠数据融合协议 SELDA[12]和基于相互监督机制的数据融合协议 WDA[13]。SELDA 协议中，传感器节点利用监视机制探测邻居节点的可用性、感知能力和路由能力而建立信任网，节点在信任网中选择一条路径传输数据。WDA 协议基于分簇结构，并且利用同级节点间的相互监督机制实现对簇头节点的安全监控。

由于 CPDA 算法的效率较高，因此，很多文献在对 CPDA 方法改进的同时提出新方法。文献[14]加入数据完整性保护机制，提出 iCPDA 算法。该算法仍然存在 CPDA 算法的缺点，尤其在增加完整性保护机制后算法复杂度增加，通信开销增大。文献[15]提出一种轻量级安全数据融合方法 LSDA，该方法采用分片重组技术，根据网络通信量大小调整簇内节点的分片数，具有较好的自适应性。实验证明，在提供相同的隐私保护程度下，LSDA 算法比 CPDA 算法具有更少的通信和计算开销。文献[5]提出一种基于分簇的低能耗隐私保护方法(LCCPDA)，该方法可以在实现数据融合的同时保护数据安全，防止数据被窃听和篡改。实验表明，LCCPDA 比 CPDA 有更好的隐私保护性和更低的数据通信量。

LCCPDA 算法和 LSDA 算法都采用了节点分片重组技术，但前者固定分片数为 3。而后者分片数至少为 3，且随着簇内节点数量增加，分片数将增加。所以，在隐私保护度相近的情况下，LCCPDA 比 LSDA 通信开销更小。但是，LCCPDA 仍存在通信开销较大和簇头被捕获概率高的问题，这是本文着重解决的问题。

3　LCSDA：一种低能耗的隐私数据安全融合协议

LCSDA 协议包括 3 个阶段：簇的形成、簇内数据融合、簇间数据融合。其中，第 1 阶段、第 3 阶段与 LCCPDA 相同。LCSDA 主要针对 LCCPDA

第 2 阶段（簇内数据融合）做了改进。

第 2 阶段包括“簇内节点分片”、“数据混杂”和“数据融合”3 个步骤。

1) 簇内节点分片

LCSDA 跟 LCCPDA 一样，都将簇内数据分成 3 片[4]。簇形成后，每个簇都包含簇头节点和簇内成员节点。假设分簇后的一个簇内包含 6 个节点（1~6），其中节点 3 为簇头节点，节点 1~6 的数据分别记为 a~f。各节点将数据分成 3 片，一片数据留在节点内，另外 2 片数据分别发送给它的 2 个邻居节点。

2) 数据混杂

LCCPDA 在数据混杂过程中，各节点随机地将数据分片传给其他节点。LCSDA 协议根据最短路径原则，将另外 2 个分片传给离自己最近的 2 个邻居节点。

假定 6 个节点组成一个簇内网络，节点间距离如图 1 所示。

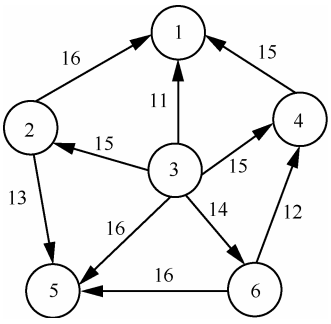

图 1 簇内网络拓扑

例如，节点 1 将数据片 a_1 留在节点内部，数据片 a_2、a_3 根据最短路径原则发送给最近的 2 个邻居节点，即节点 3 和节点 4。其他节点进行相同操作后，簇内各节点数据混杂结果如图 2 所示。

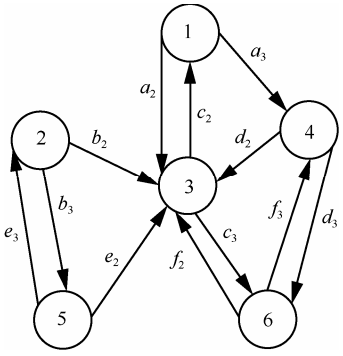

图 2 簇内各节点数据混杂

假设节点 j 经过数据混杂后的信息为 I_j，那么，每个节点经过簇内数据混杂后收到的数据

节点 1：$I_1 = a_1 + c_2$；

节点 2：$I_2 = b_1 + e_3$；

节点 3：$I_3 = c_1 + a_2 + b_2 + d_2 + e_2 + f_2$；

节点 4：$I_4 = d_1 + a_3 + f_3$；

节点 5：$I_5 = e_1 + b_3$；

节点 6：$I_6 = f_1 + c_3 + d_3$。

3) 数据融合

LCCPDA 方法在进行簇内数据融合过程中，各节点将数据直接发送给簇头，簇头节点被识别的概率高。LCSDA 方法采用最小生成树方法进行簇内数据融合。由于网络中传感器节点数量较多，呈现出边稠密特性。因此本文采用 Prim 最小生成树算法进行簇内数据融合，簇头节点不与所有节点通信，所以簇头被识别的概率降低。

基于 Prim 算法的簇内数据融合方法如下。

在每轮采集周期中，簇内节点间的拓扑用一个连通图 $N=(V, \{E\})$来表示。其中 V 表示簇内所有节点，E 表示簇内节点间边的集合。设 TE 是 N 上最小生成树中边的集合。利用 Prim 算法构造最小生成树的步骤如下。

1) 在图 $N=(V,\{E\})$中，图中每个节点自成一个连通分量。

2) 计算 $N=(V,\{E\})$中所有边的值（值等于通信距离）。

3) 在所有 $u\in U$，$v\in V-U$ 的边中，选择有最小值的边(u, v)，若该边依附的节点落在 T 中不同的连通分量上，则将此边加入到 T 中，否则舍去该边而选择下一条数值最小的边。成功加入 T 的边，则将节点 v 加到集合 U 当中，将边(u, v)加入集合 TE 中。

4) 重复步骤 3)，直到 $U=V$ 为止。此时集合 T 中所有边构成一棵最小生成树。

所有簇内节点按照以上步骤产生最小生成树后，各节点将数据发送给各自的父节点，最终将数据转发给簇头节点。

基于 Prim 算法的簇内数据融合路径如图 3 所示。

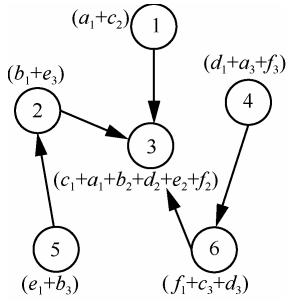

图 3 基于 Prim 算法的簇内数据融合路径

各节点分片后，通过数据混杂和数据融合后，在簇头节点 3 处得到的融合数据为

$$I = a + b + c + d + e + f$$

4　性能分析

4.1　能耗分析

假设一个簇内的节点数目为 n，簇内数据融合过程中节点间通信次数均为 $n-1$。在不考虑其他因素的条件下，能耗 E 与通信距离 d 间关系为：$E = Kd^n$。其中，K 和 n 都是常数。因此，在簇内通信次数相同时，通信距离越远，节点能耗越大。

假定簇内节点数为 6，LCCDA 协议要求各节点将数据直接传给簇头，如图 4 所示。

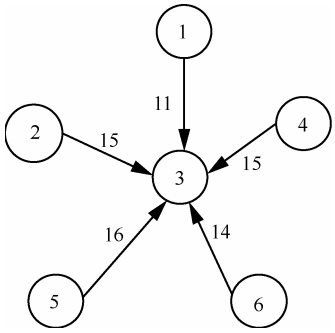

图 4　LCCPDA 协议簇内融合

LCCDA 协议通信距离和为

$$d_1 = 11 + 15 + 15 + 14 + 16 = 71$$

LCSDA 协议通过构造的簇内最小生成树路径传递数据，如图 5 所示。

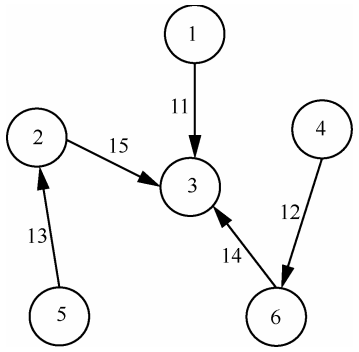

图 5　LCSDA 协议簇内融合

LCSDA 协议簇内通信距离和为

$$d_2 = 11 + 15 + 13 + 14 + 12 = 65$$

根据公式 $E = Kd^n$，且设 K=2，n=3。推出 E_1=715 822，E_2=549 250。由此可见，LCSDA 协议通信量 E_2 比 LCCPDA 协议的通信量 E_1 小。当簇内节点增多时，LCSDA 协议优势更明显。

4.2　数据通信量分析

针对一个簇内所有节点的数据通信量，对 LCSDA、LCCPDA 和 CPDA 进行比较。CPDA 协议中每个簇内所有节点的数据通信量 C_{CPDA} 为[4]

$$C_{\mathrm{CPDA}} = n^2 + n - 1 \tag{1}$$

其中，n 为一个簇内的节点数量。

LCCPDA 协议中每个簇内所有节点的数据通信量 C_{LCCPDA} 为[4]

$$C_{\mathrm{LCCPDA}} = 4n - 2 \tag{2}$$

其中，n 为一个簇内的节点数量。

LCSDA 协议中，簇内数据融合过程中节点间通信包括 2 部分：1) n 个节点分别发送分片数据给其 2 个邻节点；2) 除簇头节点外其余 $n-1$ 个节点通过最小生成树路径，将信息转发到簇头节点。因此，LCSDA 协议中一个簇内所有节点的数据通信量 C_{LCSDA} 为

$$C_{\mathrm{LCSDA}} = 2n + n - 1 = 3n - 1 \tag{3}$$

3 个算法的数据通信量如图 6 所示，从图中可以看出，LCSDA 协议的数据通信量始终小于 LCCPDA 和 CPDA，当 n 越大时，LCSDA 协议的优越性更明显。

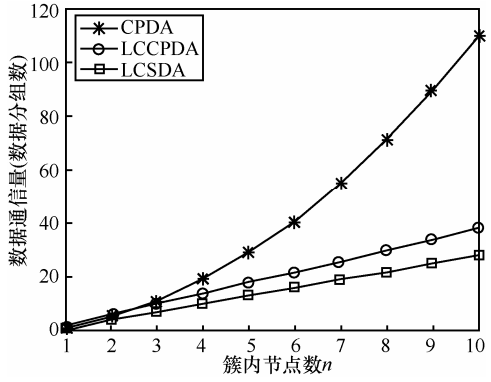

图 6　3 个算法数据通信量对比

5　结束语

本文提出一种低能耗的隐私数据安全融合方法 LCSDA，该方法根据最短路径原则为簇内每个节点选择 2 个邻居节点，并且采用 Prim 最小生成树算法建立簇内数据的融合路径。实验表明，与 CPDA 和 LCCPDA 相比，LCSDA 方法可以有效减少节点间数据通信量，同时降低节点能耗和簇头节点被捕获的概率。

参考文献：

[1] LIU C X, LIU Y, ZHANG Z J. High energy-efficient and privacy-preserving secure data aggregation for wireless sensor networks[J]. International Journal of Communication Systems, 2013, 26(3): 380-394.

[2] 杨庚,李森,陈正宇.传感器网络中面向隐私保护的高精确度数据融合算法[J].计算机学报, 2013, 36(1) :189-200.
YANG G, LI S, CHEN Z Y. High-accuracy and privacy-preserving oriented data aggregation algorithm in sensor networks[J]. Chinese Journal of Computers, 2013, 36(1):189-200.

[3] 王沁,李翀,万亚东.实时管理约束下节点级低功耗数据融合技术[J].通信学报,2008,29(11) :220-226.
WANG Q, LI C, WAN Y D. Low power data fusion for sensor node constrained by real time management in WSN[J]. Journal on Communications, 2008, 29(11):220-226.

[4] HE W B, LIV X, NGNYEN H. PDA: privacy-preserving data aggregation in wireless sensor networks[A]. Proc of the 26th IEEE International Conference on Computer Communications[C]. 2007.2045-2053.

[5] 冯艳芬, 刘宴兵. 基于分簇的低能耗数据融合隐私保护协议[J].计算机应用研究, 2013, 30(3):885-888.
FENG Y F, LIU Y B. Low energy-consuming cluster-based private data aggregation[J]. Application Research of Computers, 2013, 30(3):885-888.

[6] CAM H, OZDEMIR S, *et al*. ESPDA: energy efficient and secure pattern-based data aggregation for wireless sensor networks[A]. Proc of IEEE Sensors[C]. 2003.732-736.

[7] SANLI H O, OZDEMIR S, CAM H. SRDA: secure reference-based data aggregation protocol for wireless sensor networks[A]. Proc of the IEEE VTC Fall Conference[C]. 2004.4650-4654.

[8] WESTHOFF D, GIRAO J, ACHARYA M. Concealed data aggregation for reverse multicast traffic in sensor networks: encryption key distribution and routing adaptation[J]. IEEE Transaction on Mobile Computing, 2006, 5(10):1417-1431.

[9] CASTELLUCCIA C, MYKLETUN E, TSUDIK G. Efficient aggregation of encrypted data in wireless sensor networks[A]. Proc of Second Conference on Mobile and Ubiquitous Systems[C]. 2005.109-117.

[10] LI H, LIN K, LI K. Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks[J]. Computer Communication, 2011, 34:591-597.

[11] 杨庚, 王安琪, 陈正宇. 一种低能耗的数据融合隐私保护算法[J].计算机学报, 2011, 34(5):792-800.
YANG G, WANG A Q, CHENG Z Y. An energy-saving privacy- preserving data aggregation algorithm[J]. Chinese Journal of Computers, 2011, 34(5):792-800.

[12] OZDEMIR S. Secure and reliable data aggregation for wireless sensor networks[A]. Proc of the 4th International Conference on Ubiquitous Computing Systems[C]. 2007.102-109.

[13] DU W L, DENG J, YUNGHSIANG S, *et al*. A witness-based approach for data fusion assurance in wireless sensor networks[A]. Proc of IEEE Global Communications Conference (GLOBECOM)[C]. 2003.1435-1439.

[14] 许建, 杨庚, 陈正宇等. WSN 数据融合中的隐私保护技术研究[J].计算机工程, 2012, 38(15):134-138.
XU J, YANG G, CHEN Z Y, *et al*. Research of privacy-preserving technology in wireless sensor network data aggregation[J]. Computer Engineering, 2012, 38(15):134-138.

[15] 李万雷.无线传感器网络轻量级安全数据融合方案的研究与实现[D]. 江苏: 南京邮电大学, 2012.
LI W L. Research and Implementation on Lightweight Secure Data Aggregation Scheme in Wireless Sensor Networks[D]. Jiangsu: Nanjing University of Posts and Telecommunications, 2012.

作者简介：

谷勇浩（1980-），男，山西太原人，北京邮电大学讲师，主要研究方向为网络及信息安全等。

郭达（1976-），男，江西南昌人，北京邮电大学博士后，主要研究方向为物联网、移动互联网等。

林九川（1980-），男，江苏盐城人，公安部第三研究所助理研究员，主要研究方向为信息安全。