doi:10.3969/j.issn.1000-436x.2014.z2.036

基于隐私同态数据融合的完整性验证协议

李星，李春彦，王良民
（江苏大学 计算机科学与通信工程学院，江苏 镇江 212013)

摘 要：在无线传感器网络中的安全数据融合能够有效防止隐私泄露和数据篡改等问题，并实现高效的数据传输。由此提出一种基于隐私同态数据融合的完整性验证协议 IV-PHDA。该协议采用同态加密保证数据隐私性；利用随机检测节点对节点聚合结果的完整性进行检测，以验证聚合节点是否忠实地传输每个数据分组。通过理论分析和仿真对比，对其算法的性能进行验证，结果表明，该协议能够在网络传输的过程中检测数据的完整性，并且实现较好的隐私保护和较高的数据精确度。

关键词：数据聚合；同态加密；检测节点；数据完整性
中图分类号：TP393 **文献标识码**：A **文章编号**：1000-436X(2014)Z2-0256-05

Integrity verification protocol based on privacy homomorphism data aggregation

LI Xing, LI Chun-yan, WANG Liang-min
(School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract: Secure data aggregation in wireless sensor networks can effectively prevent privacy leakage, data tampering and other issues, it also achieves efficient data transmission. An integrity verification protocol based on privacy homomorphism data aggregation (IV-PHDA) was presented. The protocol uses homomorphism encryption to ensure data privacy, then it used random detection nodes to detect the integrity of the aggregated results to verify whether the aggregation node faithfully transmit each data packet. The theoretical analysis and simulation comparison verify the performance of the algorithm, the results show that this protocol can detect data integrity through transmission, and achieve better privacy protection and higher data accuracy.

Key words: data aggregation; homomorphism encryption; detection node; data integrity

1 引言

无线传感器网络（WSN, wireless sensor networks）具有能量受限和以数据为中心的特点，数据聚合技术通过将不同节点的数据进行合并和汇总，能够实现更高效地利用网络资源和节点能量，提高数据收集的效率和准确性[1]。在实际应用中，传感器节点所处环境可能遭受内部攻击和外部攻击等安全问题[2]，包括信道窃听和数据篡改等。因此，在设计数据聚合策略时，应充分考虑数据的隐私性和完整性等基本特征。

通常采用加密来提高数据的隐私性。Jia Guo 等人[3]给出了 WSN 中安全数据聚合的综述，将 WSN 安全数据聚合的加密方式分为逐跳加密和端到端加密，并对其具体聚合过程进行描述。在 SIA[4] 协议中，首先建立 Merkle-hash 树，采用聚合—提交—证明 3 个步骤保证聚合结果的完整性和可验证性，但该方案采用逐跳加密方式，聚合节点需要对数据进行加解密处理。He 等人[5]提出了基于数据切片的隐私保护算法 SMART，但其对数据丢失敏感。

收稿日期：2014-07-03
基金项目：国家自然科学基金资助项目（61272074）；江苏省六大人才高峰计划基金资助项目；镇江市工业科技支撑计划基金资助项目（GY2013030）
Foundation Items: The National Natural Science Foundation of China (61272074); The Six Talents Funded Project of Jiangsu Province; The Industrial Science and Technology Foundation of Zhenjiang City (GY2013030)

Westhoff 等人[6]提出了基于同态加密的数据融合隐私保护方法，但该机制不具备数据完整性检测。Chen 等人[7]提出了一种基于数据恢复的同态加密聚合隐私保护算法 RCDA，该算法能够保证数据隐私性和完整性，但其聚合延迟和通信开销较大。

针对上述问题，本文提出了一种基于隐私同态数据融合的完整性验证协议 IV-PHDA。该协议采用同态加密保证数据隐私性，利用随机检测节点对节点聚合结果的完整性进行检测，以验证聚合节点是否忠实地传输每个数据分组。结果表明，该协议能够在实现数据隐私保护和完整性检测的同时，实现高效的数据聚合。

2　系统模型

2.1　网络模型

无线传感器网络包含 3 种节点：基站 BS（base station）、聚合节点 A（aggregation node）和源节点 S（source node）。网络结构如图 1 所示，假设在本网络模型中只有一个基站，每个节点都可以对数据进行计算、发送和接收。构建以基站 BS 为根的树型结构：BS 发送聚合请求并最终得到聚合结果；聚合节点收集其子节点发送的数据，与自身数据进行融合后发送给其父节点；源节点采集监测环境中的数据并发送给其父节点。

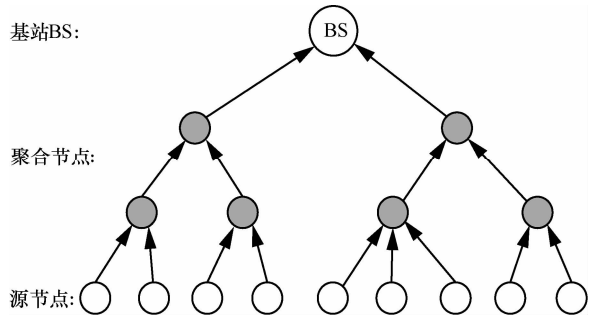


图 1　聚合树网络结构

2.2　密钥分配

本文采用随机密钥分配机制[8]，该机制的基本思想是，所有节点均从一个大密钥池中随机选取若干个密钥组成密钥链，密钥链之间拥有相同密钥的相邻节点能够建立安全链接。随机密钥分配机制的形成有 3 个阶段：密钥预分配阶段、共享密钥发现阶段和路径密钥建立阶段。

1) 密钥预分配阶段：首先产生一个大的拥有 K 个密钥的密钥池和密钥标识，然后随机选取不同的 k 个密钥组成密钥链，再把不同的密钥链分配给不同的节点。

2) 共享密钥发现阶段：每个节点都要发现周围与其有共享密钥的节点，只有存在共享密钥的节点之间才被认为是连接的。

3) 路径密钥建立阶段：若 2 个节点之间没有共享密钥，则可通过存在共享密钥的路径来建立链路密钥。

任意 2 个节点能够共享一个相同的密钥概率为

$$P_{\text{connect}} = 1 - \frac{\left((K-k)\ !\right)^2}{(K-2k)\ !K\ !} \tag{1}$$

在同样的网络中，攻击者作为节点，也采用相同的随机密钥分配机制，也需要从这个拥有 K 个密钥的密钥池中随机选取 k 个密钥，而攻击者获得该密钥的概率为

$$P_{\text{compromised}} = \frac{k}{K} \tag{2}$$

在通常情况下，$P_{\text{compromised}}$ 是一个很小的数。

3　IV-PHDA：基于隐私同态数据融合的完整性验证协议

本文提出了一种基于隐私同态数据融合的完整性验证协议 IV-PHDA，该协议的具体内容大致可以分为 4 个部分：系统初始化、数据加解密、数据聚合、数据检测。

3.1　系统初始化

3.1.1　构建聚合树

构建聚合树的基本过程如下：BS 首先将查询请求发送给较近的某一节点（假设为 N_0），N_0 节点通过逐跳的方式将该查询请求广播到整个无线传感器网络。当网络中的所有节点都接收到这个查询请求时，一个以 BS 为根的数据聚合树也就生成了，聚合树的具体构建方法同 TAG[9]。

3.1.2　检测点的选取

为了对数据在传输过程中的完整性进行检测，本文通过选取检测点对节点数据的完整性进行检测，以验证聚合节点是否忠实地传输每个数据分组。

若事先对网络中的检测点名单进行固定，当一定数量的检测点被攻击时，攻击者的攻击行为将无法被发现，这种方法并不安全。因此，本文采用一种随机的检测点选取方法（DNS, detection nodes selection），以降低检测节点被攻击的风险，增强网

络的安全性。这个选取方法主要包括 2 个步骤：节点初始化和检测节点确认。

1) 节点初始化。在选取检测点前，每个节点加载 2 个函数：一个单向散列函数 F(*ID*, *x*)和一个映射函数 $f_p(y)$。其中，*ID* 为传感器节点的 *ID*，*p* 是一个预先定义的概率值，函数 $f_p(y)$的值域为(0, 1)，并且当 *y* 在 F 函数的值域内时，f_p 的值以概率 *p* 映射到 1，以概率(1−*p*)映射到 0。

2) 检测节点确认。源节点为每个数据生成一个随机数 *r*，这个数据经过逐跳传送给基站。传输过程中的聚合节点在接收到这个数据时，对 $f_p(F(ID, r))$的值进行检查，若该值等于 1，则该节点被选为检测点，并生成一个确认数据向源节点方向传递。当一个子节点收到其父节点发送的确认数据时，对该数据进行验证：通过计算 $f_pF(ID', r))$是否为 1 来判断该数据是否来自一个真实的检测点（其中，*r* 为随机数，*ID'*是检测点的 *ID*）；通过该数据的消息认证码来验证该数据的内容是否被篡改。

算法的流程如图 2 所示。这种检测点选取方法的优势在于其随机性和动态性。检测节点选取的随机性使攻击者无法确定具体的检测节点，每个聚合节点都可能成为检测节点；而检测节点选取的动态性使检测节点的名单不确定，攻击者无法预知下一次的检测节点。

由于检测节点是随机选取的，因此可能存在网络中无节点被选为检测节点的情况，这时整个网络只有基站是检测点，即变为端到端检测。这种情况不会对检测节点选取方案的正确运行产生影响。

3.2 数据加解密

本文采用同态加密技术[2]对源节点从监测环境中采集到的数据进行加密处理。其中，公共参数 *d*、整数 *g* 和密钥 *k*=(*r*, *g*')为所有节点已知。加解密方法如下。

加密：将明文随机分成 *d* 份密文，加密公式为

$$Ek(a)=(a_1r^1 \bmod g,\ a_2r^2 \bmod g,\ \cdots,\ a_dr^d \bmod g) \quad (3)$$

解密：通过 $r^{-i} \bmod g$ 计算第 *i* 个坐标的标量积来恢复 $a_i \bmod g$，*a* 的计算式为

$$D_k\left(E_k\left(a\right)\right)=\left(\sum_{i=1}^{d}a_i\right) \bmod g' \quad (4)$$

3.3 数据聚合

由同态加密的性质[10]可知，经过同态加密的数据可在聚合节点处直接进行聚合处理，聚合节点不需对数据进行加解密。

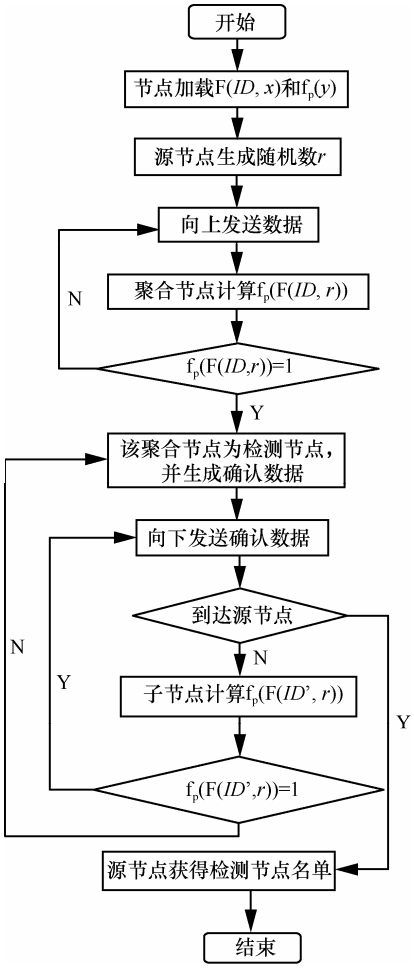


图 2 检测点选取流程

数据聚合包括求和、求平均值、计数、最大值和最小值聚合等，由于其他聚合函数都可以通过转化成求和函数计算[11]，因此本文只考虑求和的数据聚合函数，记为

$$y(t)=\sum_{i=1}^{N}d_i\,(t) \quad (5)$$

3.4 数据检测

采用同态加密的方法可保证数据在网络传输过程中的端到端隐私性，但不能保证数据在传输过程中的正确性。IV-PHDA 协议的数据检测部分除了在基站对最后还原的聚合结果进行检测外，还利用网络中随机选取的检测点对网络中聚合节点的聚合结果进行验证，以实现更高效的数据传输。

综合上述 4 个部分，本文提出了一种可检测完整性的数据融合算法 IDDA，在源节点处利用同态加密方法对源节点数据进行加密保证隐私性；在数

据传输过程中利用网络中随机选取的检测点对聚合节点的聚合计算结果进行检测；在基站对还原的聚合结果进行验证。整个算法具体过程如图 3 所示。

Algorithm. IDDA
1) Source nodes encrypt the data.
2) **for** $1 \leqslant i \leqslant l$ (l is the number of the source nodes), **do**
3) $s_i = \mathrm{E}k(S_i)$
4) Source nodes transmit s_i to Aggregation nodes
5) Aggregation nodes aggregate the data
6) **for** $1 \leqslant i \leqslant l$ (l is the number of the child nodes of one aggregation), **do**
7) $a = \sum s_i$
8) Check if there has some detection nodes
9) **if** aggregation node A_j is a detection node
10) **then** the next aggregation node receive the data from its child nodes (A_j is included)
11) the next aggregation node's other child nodes send their data to A_j
12) the next aggregation node and A_j aggregate the data, check the aggregated results
13) **if** the results are not equal **then**
14) drop the data and inform it
15) **else** the next aggregation node transmit the data to next node
16) **else** the next aggregation node sends the aggregated data to next node
17) Repeat step 8 to 16 until the data is transmitted to BS
18) BS decrypts the data it receives, and verifies the result

图 3　IDDA 算法

4　性能分析和仿真结果

本文主要从隐私保护、数据完整性、通信开销和精确度方面对 IDDA 算法的性能进行分析。选择 NS2 平台进行模拟仿真实验，网络配置环境如下：400 m×400 m 区域中随机分布 600 个节点，节点传输距离为 50 m，背景噪声为−105.0 dBm，高斯白噪声为 4 dB，数据传输率为 1 Mbit/s。

4.1　隐私保护

WSN 中最常见的攻击是窃听攻击，攻击者通过尝试恢复截获的密文来获取信息。本文加密处理采用随机划分，即密文的结果是概率性的，所以可以抵抗已知明文攻击和已知密文攻击，能够很好地保证数据的隐私性。检测节点选取的随机性，也使检测节点名单不容易被攻击者获取，

4.2　数据完整性

数据完整性是指接收方接收到的消息与发送方传来的一致，过程中没有被攻击者篡改或伪造信息。同态加密方法能够在基站通过恢复数据（解密）的方式对聚合结果的正确性进行检测，但不具备数据传输过程中完整性检测能力。

以图 4 网络结构为例，其中 A_2 为检测节点，在 A_2 处会对 A_3 的聚合结果进行验证。若在 A_3 处接收到的 A_1 的数据被篡改为(5, 62)，此时 A_3 的聚合结果会变为(14, 2)，而检测节点的聚合结果为(14, 22)与 A_3 的聚合结果不一致，此时 A_3 丢弃该数据并告知 A_1 和 A_2。由此可知，IDDA 算法能够良好检测到数据的完整性是否破坏，提高网络在数据传输过程中完整性检测的能力。

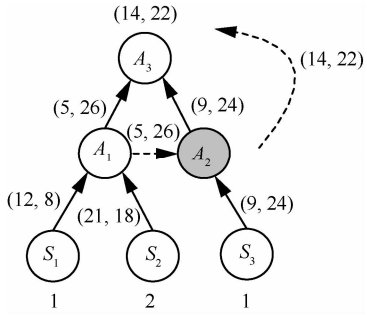


图 4　网络部分节点运行 IDDA 过程

4.3　通信开销

节点的通信开销与节点能量消耗直接相关。从理论上分析，令 N 表示 WSN 中节点的数量，则 TAG[9]算法的数据通信开销为 $O(N)$，SMART[5]算法的通信开销为$(J-1)N$（J 为 SMART 算法中的分片数目），而 IDDA 算法相较于 TAG 算法多了检测节点的 2 次通信开销，因此 IDDA 算法的通信开销为 $O(N)+2k$（k 为 WSN 中检测节点的数量）。

图 5 为 TAG、SMART 和 IDDA 算法通信开销的对比，3 种算法的通信开销都随着节点的增加而增加，其结果与理论分析一致。

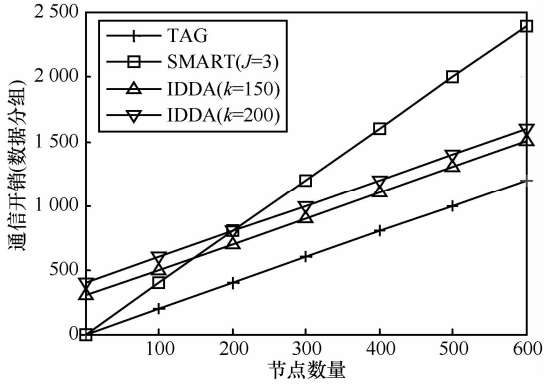


图 5　TAG、SMART 和 IDDA 算法的通信开销

4.4　精确度

定义 WSN 数据聚合的精确度为实际得到的数据聚合结果与节点采集数据之和的比值。理想情况下，在传输过程中没有数据丢失，TAG、SMART 和 IDDA 算法的数据聚合结果都应该达到 100%的

精确度。但由于网络无线通道的冲突性、节点失效性和数据处理的延迟性等导致在实际应用中传输信息的丢失，从而影响精确度。

图 6 为 TAG、SMART 和 IDDA 3 种算法的精确度对比，仿真是针对 Epoch Duration 的变化进行的。由于 IDDA 算法在数据的传输过程中对数据聚合结果进行检测，因此该算法的精确度要高于 TAG 算法，由图 6 可知其结果与理论分析一致。

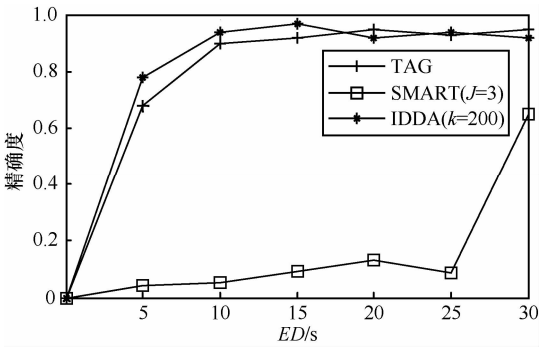


图 6　TAG、SMART 和 IDDA 算法的精确度

5　结束语

在无线传感器网络中，攻击者可以利用受损节点窃听和篡改数据信息。本文提出了一种基于隐私同态数据融合的完整性验证协议，基于同态加密实现数据隐私性保护，基于随机检测节点对网络中节点的聚合结果进行验证，从而检测出节点是否正常传送数据分组。实验结果和理论分析表明，该算法能够检测数据在网络传输过程中的完整性，并且能够在较低通信量的基础上具有较好的隐私性和较高的精确度。

参考文献：

[1] RAMESH R, VARSHNEY P K. Data aggregation techniques in sensor networks: a survey[J]. Communications Surveys & Tutorials, 2006, 8(4):48-63.

[2] OZDEMIR S, YANG X. Secure data aggregation in wireless sensor networks: a comprehensive overview[J]. Computer Networks,2009, 53 (12):2022-2037.

[3] GUO J, FANG J A, CHEN X M. Survey on secure data aggregation for wireless sensor networks[J]. IEEE Internaional Conference on SOLI, 2011, 7: 138-143.

[4] PRZYDATEK B, SONG D, PERRIG A. SIA: secure information aggregation in sensor networks[A]. Proc of 1st Conference on Embedded Networked Sensor Systems[C]. Amsterdam: IOS Press, 2003. 255-265.

[5] HE W, LIU X, NGUYEN H. PDA: privacy-preserving data aggregation in wireless sensor networks[A]. Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM)[C]. 2007.2045-2053.

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

[7] CHEN C M, LIN Y H, LIN Y C. RCDA: recoverable concealed data aggregation for data integrity in wireless sensor networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2012,23(4):727-734.

[8] ESCHENAUER L, GLIGOR V D. A key-management scheme for distributed sensor networks[A]. Proceedings of the 9th ACM Conference on Computer and Communications Security[C]. Washington, 2002.41-47.

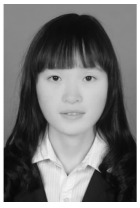
[9] MADDEN S, FRANKLIN M J, HELLERSTEIN J M. TAG: a tiny aggregation service for ad-hoc sensor networks[A]. Proceedings of the 5th symposium on Operating systems design and implementation[C]. New York, USA: ACM, 2002.131-146.

[10] DOMINGO-FERRER J. A provably secure additive and multi- plicative privacy homomorphism[A]. Proceedings of the Information Security Conference[C]. 2002.471-483.

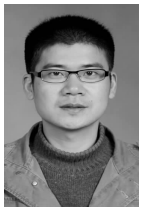
[11] CASTELLUCCIA C, MYKLETUN E, TSUDIK G. Efficient aggregation of encrypted data in wireless sensor networks[A]. The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services[C]. San Diego, CA, 2005.109-117.

作者简介：

李星（1991-），女，福建宁德人，江苏大学硕士生，主要研究方向为传感器网络安全数据融合。



李春彦（1988-），女，山东菏泽人，江苏大学硕士生，主要研究方向为车联网中的入侵检测技术。



王良民（1977-），男，安徽潜山人，江苏大学教授、硕士生导师，主要研究方向为信息安全与物联网工程。