一种低耗能的数据融合隐私保护算法

杨 庚 王安琪 陈正宇 许 建 王海勇
(南京邮电大学宽带无线通信与传感网技术教育部重点实验室 南京 210003)
(南京邮电大学计算机学院 南京 210003)

摘 要 物联网中的隐私保护是实际应用中要解决的关键问题之一,作为物联网组成部分的无线传感器网络,希望在进行精确数据融合的同时,又能保护个人的隐私.文中提出了一种新的低能耗无线传感器网络数据融合隐私保护算法 ESPART.一方面算法依靠数据融合树型结构本身的特性,减少数据通信量;另一方面算法分配随机时间片,避免碰撞.同时限制串通数据范围,降低数据丢失对精确度的影响.仿真结果显示,相比于 SMART 算法,ESPART可以在有效保护数据隐私的前提下,花费与 TAG 算法相同的时间和较少的数据通信量,得到精确的数据融合结果.

关键词 物联网;无线传感器网络;数据融合;隐私保护;低耗能;高效能
中图法分类号 TP393 **DOI 号**: 10.3724/SP.J.1016.2011.00792

An Energy-Saving Privacy-Preserving Data Aggregation Algorithm

YANG Geng WANG An-Qi CHEN Zheng-Yu XU Jian WANG Hai-Yong
(*Key Laboratory of Broadband Wireless Communication & Sensor Networks Technology of Ministry of Education, Nanjing University of Post & Telecommunications, Nanjing* 210003)
(*College of Computer Science & Technology, Nanjing University of Post & Telecommunications, Nanjing* 210003)

Abstract Privacy preserving plays an important role in application of the Internet of Things (IoT). As a part of the IoT, Wireless Sensor Networks (WSNs) should provide the privacy preserving in data aggregation. This paper presents a novel energy-saving private-preserving aggregation scheme (ESPART) for WSNs, 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 collusion data to reinforce data_loss resilience. Compared with the SMART algorithm, the simulation results show that ESPART can preserve data privacy, get accurate data aggregation results while taking the same epoch duration as TAG, and have less communication overhead.

Keywords Internet of Things; wireless sensor network; data aggregation; privacy preserving; energy saving; energy efficiency

1 引 言

隐私保护是物联网信息机密性的直接体现,如感知终端的位置信息、采集的数据信息等是物联网的重要信息资源之一,也是需要保护的敏感信息.另外在数据处理过程中同样存在隐私保护问题,如基于数据挖掘的行为分析等等.在物联网的实际应用

收稿日期:2011-01-17;最终修改稿收到日期:2011-04-06.本课题得到国家“九七三”重点基础研究发展规划“物联网混杂信息融合与决策研究”课题(2011CB302903)、国家自然科学基金项目“无线传感器网络组播广播安全关键技术研究”(60873231)和江苏省自然科学基金(BK2009426)资助.**杨 庚**,男,1961 年生,博士,教授,中国计算机学会高级会员,主要研究领域为计算机通信与网络、网络安全、分布与并行计算.E-mail: yangg@njupt.edu.cn. **王安琪**,男,1987 年生,硕士研究生,主要研究方向为无线传感器网络数据融合与安全. **陈正宇**,男,1978 年生,博士研究生,讲师,主要研究方向为信息安全、隐私保护. **许 建**,男,1980 年生,博士研究生,讲师,主要研究方向为数据融合与安全. **王海勇**,男,1979 年生,博士研究生,讲师,主要研究方向为数据融合与安全.

中，需要解决其信息隐私保护问题.

无线传感器网络是物联网的主要组成部分，它的任务主要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. 例如，健康监控系统利用无线传感器网络监测病人，收集并汇聚病人实时信息，通过分析这些数据可以得到相应的结论和处理手段.

无线传感器网络中的节点通常是资源受限的，同时由于能量资源关系到整个传感网络的生存时间，因此能量消耗通常是传感网中数据融合需要重点考虑的问题. 文献[1]表明使用 Micadot 节点发送 1bit 数据需消耗能量约 4000nJ，处理器执行一条指令需消耗的能量仅约为 5nJ，因此减少通信数量可以有效地减少能量消耗. 另一方面，在实际应用中，人们希望通过传感器网络收集其覆盖范围内的相关检测数据，但在数据的收集与传输过程中，存在大量的数据冗余与碰撞，影响了传感网有效数据的采集，降低了网络的生存时间. 因此，设计有效的数据融合算法是十分必要的[2]，它可以剔除感知信息中的冗余数据，在得到有效数据的前提下，减少数据传输量，也相应减少了数据传输过程中的碰撞，减轻网络拥塞，节省节点能量，达到延长网络生命周期的目的。文献[3]中提出的 TAG(Tiny AGgregation)算法就是一种典型的应用在无线传感器网络中的数据融合技术.

基本的数据融合技术一般不提供数据的隐私保护机制. 然而在实际应用中，隐私保护机制是不可或缺的，比如在健康监控系统中，节点可以获得病人的体温、血压、脉搏等数据，这些数据属于个人隐私，病人不想其被其它人得到. 在基本的数据融合技术中，节点使用无线信道，将数据沿着数据融合技术构造的数据融合树层层向上传递并融合，最终 QS(Query Server)得到需要的融合数据. 然而，由于无线传输的特性，节点间传输的数据易于被捕获和偷听，并且网内的可信任父节点可以得到子节点的数据，如果攻击者破解了无线链路或捕获了父节点，病人的隐私便暴露了. 无线传感器网络数据融合隐私保护技术就是在保证数据融合结果正确的情况下，即使是传输的数据被外部捕获与解密或是内部其它可信节点被捕获，也能够阻止隐私数据被获取的技术.

现有的研究[4-12]已经提出了一些隐私保护方案，它们各自有自己的适用范围，而且部分解决方案还有一些问题需要进一步解决. 例如，Girao 和 Castelluccia 提出的一种数据融合隐私保护解决方案[7-8]. 它们使用同态加密，可以让节点在不需要解密数据的条件下，对数据实施有效的融合，但该种方案对 QS 不具备隐私保护性. Conti 等人提出了一种提高鲁棒性的数据融合隐私保护算法[9]. 它加强了鲁棒性，而且其节点计算复杂度与数据传输量都不大，但需要基于簇状结构的无线传感器网络. Bista 等人提出了一组新型数据融合隐私保护解决方案[10-12]. 它们都有着计算复杂度与数据传输量小的优点，但是它们都对 QS 不具备隐私保护性.

本文提出了一种能量节省的数据融合隐私保护算法 ESPART，它对 SMART[4] (Slice-Mix-AggRegaTe)做了相应的改进，采用逐跳(hop-by-hop)的数据融合方式，具有计算复杂度与数据传输量小的优点，并对 QS 具备隐私保护性. 它采用给每个节点分配随机时间片的办法，以避免节点间的碰撞，并限制节点间串通(collusion)的数据范围，降低数据丢失对精确度的影响，使得 ESPART 在精度要求相近的情况下，需要的数据融合时间(Epoch Duration)较少. ESPART 采用依靠数据融合树形结构本身特性的办法，消除 SMART 算法中不必要的节点通信，使其在同等隐私保护安全性情况下，数据传输数量减少为 SMART 的 50%左右.

本文第 2 节对相关工作进行讨论，详细分析 SMART 算法的安全性、通信复杂性和精确度等性能；第 3 节描述本文所采用的系统模型；第 4 节给出一种新的数据融合方法以及具体的算法；第 5 节用理论方法和仿真实验，分析本文所述方法在隐私保护和能量消耗等方面的优越性，并与其它方法进行比较。

2　相关工作

Wenbo 等提出了数据融合算法 PDA(Privacy-preserving Data Aggregation)[4]，其中包含了两种算法 CPDA(Cluster-based Private Data Aggregation)和 SMART，但是 CPDA 计算量大，而 SMART 数据通信量也太大，并且对数据丢失敏感，需要两倍于 TAG 的时间来获得相对精确的结果.

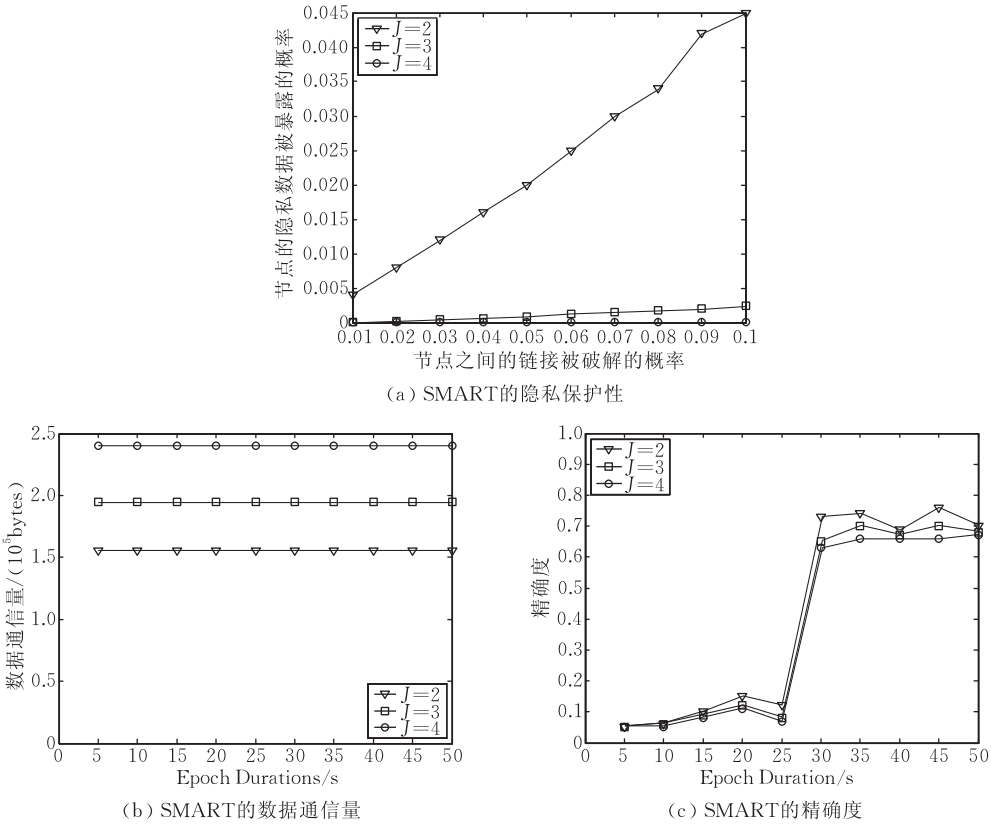
与本文提出的 ESPART 算法关系最为密切的是 SMART 算法[4]. SMART 算法使用逐跳(hop-by-hop)的数据融合方式和点到点(node-to-node)的加解密模式，它能够在保证数据融合结果精确程度的情况下，阻止外部入侵者的攻击，保证内部可信节

点以及 QS 获得隐私数据.

SMART 算法步骤分成三步,分片算法(Slicing)、混合算法(Mixing)和融合算法(Aggregating). 在 SMART 算法中,节点先将自身测得数据进行分片,然后将这些分片依次发送给随机选取的邻居节点们,最终沿数据融合树向上进行数据融合.

SMART 算法的节点计算复杂度为 $O(1)$,由于对自身测得数据进行分片并相互交换混合的原因,同时 SMART 需要传输的数据大约是 TAG 的 $(J+1)/2$ 倍(J 是 SMART 算法中的分片数目,通常选取 $J\geq 3$ 以达到隐私保护的效果),而且 SMART 对数据丢失十分敏感,以至于大约需要 2 倍于 TAG 的数据融合时间才能得到相对精确的结果[4].

图 1(a)显示了 SMART 的隐私保护性,可以看出,当 $J\geq 3$ 时,节点的隐私数据被暴露的概率小于 0.5%;图 1(b)显示了 SMART 的数据通信量,可以看出,其值随 J 的增加而增加;图 1(c)显示了 SMART 的精确度,可以看出,在 Epoch Duration=25s 前,数据融合结果基本不能体现实际的数值.



(a) SMART的隐私保护性

(b) SMART的数据通信量

(c) SMART的精确度

图 1 SMART 性能[4]

3 系统模型

在本文中,无线传感器网络由一个连通图 $G(V,E)$来表示,其中的顶点 $v(v\in V)$表示无线传感器网络中的节点,边 $e(e\in E)$表示节点间的通信链路. 记无线传感器网络中节点的数量为 $N=|V|$.

在无线传感器网络中,包含了 3 种类型的节点:QS(Query Server)节点、融合节点和叶子节点,我们只考虑网络中只有一个 QS 节点的情况. 在传统的数据融合技术中,构造的数据融合树以 QS 节点为根,得到数据融合的最终结果;融合节点融合从其子节点接收的数据和本身采集到的数据,并向上传递给其父节点;叶子节点只负责采集数据并传给其父节点.

我们定义数据融合函数为 $y(t)=f(d_1(t),d_2(t),\cdots,d_N(t))$,$d_i(t)(i=1,2,\cdots,N)$表示节点 i 在 t 时刻采集到的数据(如图 2 所示). 由于很多典型的数据融合函数,比如 count、average、max、min 等都可以化简为 sum 函数[8],因此在本文中我们以 sum 函数为研究对象,记 $y(t)=\sum_{i=1}^{N}d_i(t)$.

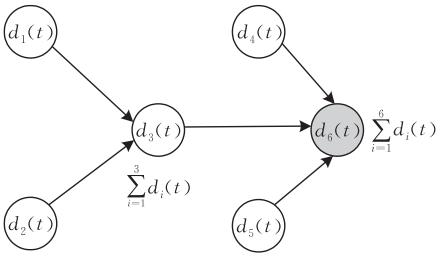


图 2　数据融合 sum 函数示意图

无线传感器网络在很多实际应用中都必须考虑隐私保护，以保证每个节点采集到的数据只能由授权的用户访问. 对数据的隐私保护的手段之一是加密. ESPART 算法采用逐跳的数据融合方式和节点到节点的加解密模式，其密钥分配与管理策略与 SMART 算法一样，采用随机密钥分配机制[13]，即首先生成一个拥有 K 个密钥的密钥池，每个节点从这个密钥池中随机选取 k 个密钥，如果邻居节点与本节共享一个相同的密钥，记为 key_s，此时这两个相邻节点间就可以通过这个共享密钥 key_s 建立起一条安全链接. 任意两个节点能够共享一个相同的密钥的概率为 $p_{\text{connect}}=1-\dfrac{((K-k)!)^2}{(K-2k)!K!}$.

对无法找到共享密钥对的相邻节点可以通过建立一条多跳的链接，实现节点间的加密解密过程[13]. 同样，在此无线传感器网络中，作为偷听者的第三方节点，采用相同的随机密钥分配机制，即也从这个拥有 K 个密钥的密钥池中随机选取 k 个密钥，那么如果在这 k 个密钥中包含有 key_s，则此偷听节点可以使用这个密钥 key_s 对已加密信息进行解密，此时先前建立的一条安全链接便不再安全. 偷听节点获得 key_s 的概率为 $p_{\text{overhear}}=\dfrac{k}{K}$.

假定密钥池中共包含 $K=10000$ 个密钥，每个节点随机选取 $k=50$ 个密钥，那么任意两个节点间可以共享相同密钥的概率为 $p_{\text{connect}}=78\%$；一旦一对节点选取一个共享密钥 key_s 建立起一条安全链接，那么偷听节点拥有共享密钥 key_s 的概率为 $p_{\text{overhear}}=\dfrac{k}{K}=0.5\%$. 在通常情况下，$p_{\text{overhear}}$ 是一个很小的数.

4　ESPART：一种低耗能的数据融合隐私保护算法

设实际测量数据值的下界为 $V_{\text{lowerbound}}$，上界为 $V_{\text{lowerbound}}$，节点产生的随机数表示为 $r(r\in(-1,1))$，其在 $(-1,+1)$ 区间内均匀分布. 同时为了保证隐私保护性，设节点需要的最小出入度度数为 $MinDeg$；为得到可靠的数据融合结果，设数据融合循环次数为 $Loop$，每个节点分配 1byte 空间用以记录本节点的循环次数 $loop$. QS 为一个时延 $EpochD$(Epoch Duration)的时间，节点在这个 $EpochD$ 时间内分配一个串通时延 $CollD$(Collusion Duration)和一个融合时延 $AggD$(Aggregation Duration)，所有节点在 $CollD$ 时间内进行串通，在 $AggD$ 时间内向上传输并融合.

$$CollD=\frac{EpochD\cdot MinDeg}{MinDeg+MAX\_DEPTH},$$

$$AggD=\frac{EpochD\cdot MAX\_DEPTH}{MinDeg+MAX\_DEPTH},$$

其中 MAX_DEPTH 表示此无线传感器网络中数据融合树的最大深度.

算法 1.　ESPART 算法.

1. 准备阶段.

1.1. 计算节点集 S_i：每一节点 i $(i=1,2,\cdots,N)$ 随机在 h 跳内选取一个节点集 S_i $(|S_i|=MinDeg)$，对于稠密的无线传感器网络，h 值为 1 即可.

1.2. 计算融合树与参数：在一个 $EpochD$ 时间内，使用 TAG 算法，产生数据融合树，在每一节点 $i(i=1,2,\cdots,N)$ 向上进行数据融合时，记录其子节点的数目，记为参数 deg_i. 明显有叶子节点的 $deg_i=0$. 此时得到一个结果，并且节点将循环次数 $loop$ 加 1.

2. 节点串通通信阶段.

2.1. 计算时间片：由 $MinDeg$ 的大小确定分配的时间片的多少，记 t_p $(p=1,2,\cdots,MinDeg)$ 为在 $CollD$ 内分配的时间片，它们在 $CollD$ 内均匀分布，则有 $t_p=\dfrac{CollD}{MinDeg}$ $(p=1,2,\cdots,MinDeg)$.

2.2. 串通节点间进行数据通信：

首先网络节点处于在 $CollD$ 内分配的时间片 t_1 内，让节点的 deg_i 为 0 的节点们，即叶子节点们在分配的时间片 t_1 中随机选择时间与其它节点进行串通；接着在分配的时间片 t_2 内，让节点 deg_i 为 1 的节点们在分配的时间片 t_2 中随机选择时间与其它节点进行串通，一直到让节点 deg_i 为 $MinDeg-1$ 的节点在其分配的时间片 t_{MinDeg} 内与其它节点进行串通. 具体为：

① 节点 i 在此次循环时实际测得数据记为 d_i，伪数据记为 f_i，节点首先验证自己的 deg_i 是否小于 $MinDeg$. 若非，则不需要请求发送串通数据，在 $CollD$ 内只需接收即可，$CollD$ 结束后，直接进入步 3；若是，则在 S_i 中顺序选择节点 j，向节点 j 发出种子 $Seed_{ij}$，并将 deg_i 加 1，从节点自身记录的数据(发出或接收第一个种子前，其值为 d_i)中减去 $Seed_{ij}$. 种子的计算公式为

$$Seed_{ij}=\frac{r\cdot(V_{\text{uperbound}}-V_{\text{lowerbound}})}{MinDeg}.$$

此时节点 i 中的伪数据为 $f_i=d_i-Seed_{ij}$.

若节点 i 收到加密的种子,则使用与邻居节点共享的密钥解密,得到种子 $Seed_{ij}$,然后将得到的种子与自身记录的数据相加,并将 deg_i 加 1,此时得到的伪数据为 f_i+Seed_{ji}.

② 跳到步①开始执行.

3. 数据融合阶段.

在 $AggD$ 时间内,依步 1 中建立的数据融合树,使用 TAG 算法,从底层节点向上层节点逐步进行数据融合,最终在 QS 得到数据融合结果,此时节点将循环次数 $loop$ 加 1.

4. 算法循环与结束阶段.

若 $loop$ 小于 $Loop$ 则跳到步 2 开始执行,否则算法结束.

图 3 显示了在 $N=8$ 的无线传感器网络中,并在给定 $Loop=2$,$MinDeg=3$,$h=1$ 的条件下,执行一次 ESPART 算法的具体步骤. 算法中,相邻节点间的每一次数据传输都需要使用随机密钥分配机制[13]生成的密钥进行加密与解密.

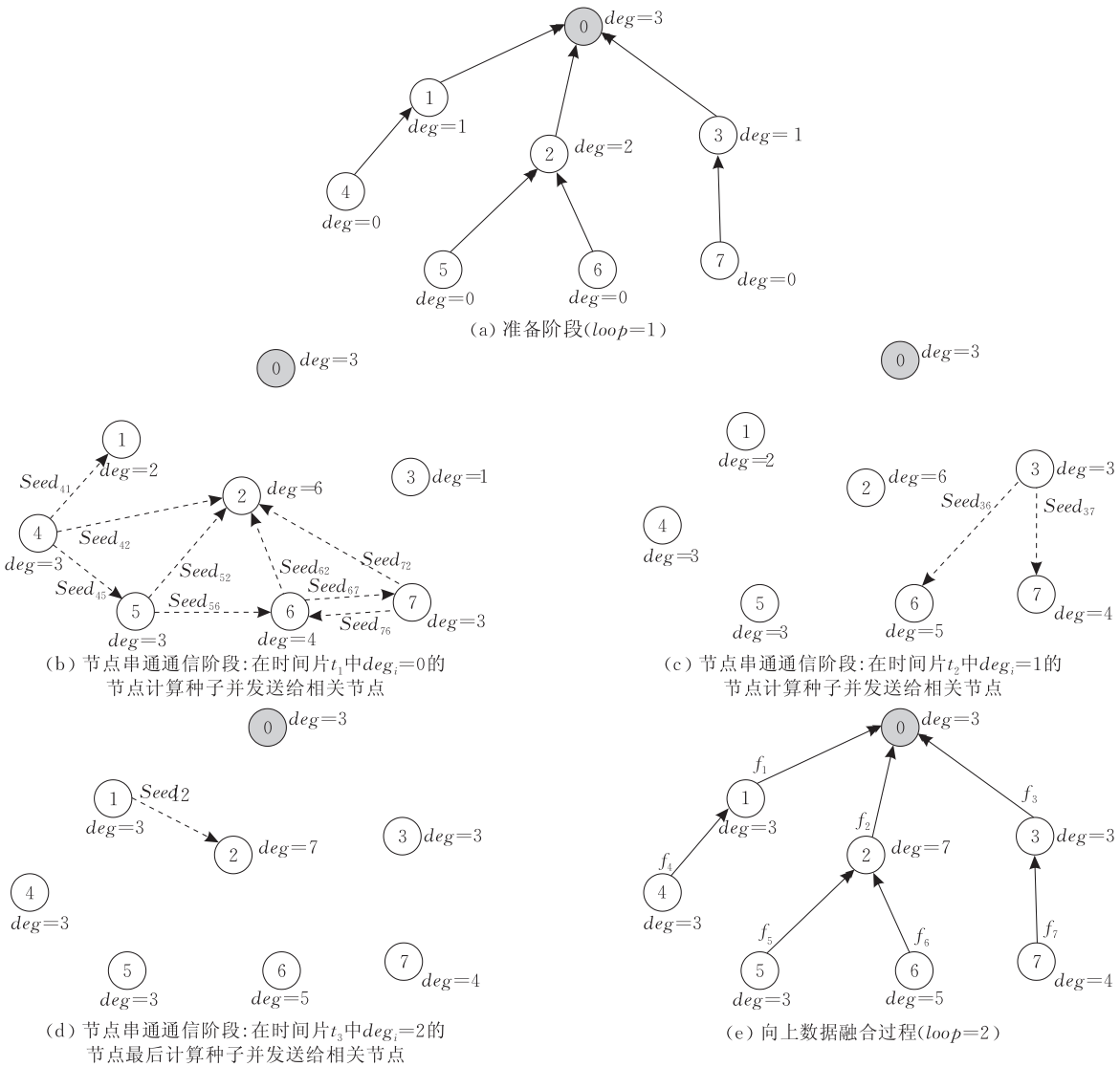


图 3 ESPART 算法示意图

图 3(a)表示处于准备阶段的无线传感器网络中各节点的情况,在此阶段中节点除了做 TAG 算法,向上融合数据得到结果外,还要记录其子节点的个数并保存于 deg 变量中,并将 $loop$ 加 1.

由于 $MinDeg$ 定为 3,故将节点间串通通信阶段分为 $MinDeg=3$ 个时间片,记为 t_1、t_2、t_3,它们依次进行. 图 3(b)~(d)分别表示网络中的节点依次处于串通通信阶段时间片 t_1、t_2、t_3 中,网络中 deg 变量为 0、1、2 的节点,向邻居节点们发出种子 $Seed$ 的过程. 节点每发一个 $Seed$ 或者每收到一个 $Seed$,就将其自身的 deg 变量加 1,直到 deg 不小于 $MinDeg$. 通过约定的串通时间分片通信模式发送数据,可以

避免所有节点同时刻发送，减少网络中的碰撞次数. 并且通过确定自身 deg 变量的方法来决定是否需要发送数据，可以减少网络中的数据传输数量. 并且每个节点发送的 $Seed$ 都经过计算控制在一定范围内，可以减小数据丢失带来的影响.

图 3(e)表示网络中节点依照准备阶段建立好的数据融合树向上融合，最终在 QS 得到数据融合结果，并将 $loop$ 加 1. 此时 $loop$ 已经等于 $Loop$，算法结束.

5　性能分析

在本节中我们主要从隐私保护性、数据通信量、精确度这三个方面分析 ESPART 的性能. TAG[3] 是应用在无线传感器网络中的一种典型数据融合技术，它不提供隐私保护功能，我们使用它作为 ESPART 的数据通信量、精确度的对比项.

本文采用 TOSSIM[14] 软件进行 TAG 和 ESPART 算法的仿真，具体的网络环境配置为：600 个节点随机分布于 400m×400m 区域中，无线信道对称，标准室内环境，背景噪声为 −105.0dBm，高斯白噪声为 4dB，节点的数据传输速率为 1Mbps，节点的灵敏度为 −108.0dBm，传输距离为 50m.

5.1　隐私保护性

对于一个节点来说，记其测得的数据被获取的概率为 $P(q)$，q 表示节点之间的链接被破解的概率，$q\approx p_{\text{overhear}}$.

在 ESPART 算法中，每个节点的度数至少是 $MinDeg$. 除非将节点的入度与出度链接都破解，节点的真实数据才会暴露，因此

$$P(q)=\sum_{k=MinDeg}^{d_{\max}} P(deg=k)\cdot q^k \qquad (1)$$

其中，$d_{\max}$ 表示网络中节点 deg 的最大值，$P(deg=k)$ 表示网络中节点的度数为 k 的概率. 记网络中度数为 k 的节点数量为 $N(deg=k)$，则 $P(deg=k)=\dfrac{N(deg=k)}{N}$.

相应的在 SMART[4] 中，除非节点所有的 $J-1$ 个出度与所有的入度链接都被破解，节点的真实数据才会暴露，此时

$$P(q)=q^{J-1}\sum_{k=0}^{d_{\max}} P(in\_degree=k)\cdot q^k \qquad (2)$$

其中，$d_{\max}$ 表示网络中节点入度的最大值，$P(in\_degree=k)$ 表示网络中节点入度为 k 的概率. 记网络中入度为 k 的节点数量为 $N(in\_degree=k)$，则 $P(in\_degree=k)=\dfrac{N(in\_degree=k)}{N}$.

对式(1)和(2)进行对比分析可知，为了保证节点的隐私数据被暴露的概率足够小，ESPART 算法与 SMART 算法不同，它并不保证节点的出度数量 $J-1$，它保证的是节点的出入度之和 $MinDeg$. 无线传感器网络隐私保护性强度与节点的隐私数据被暴露的概率成反比. 理论上，在 $MinDeg$ 与 $J-1$ 相等的情况下，由于 ESPART 算法串通的次数小于 SMART 算法串通的次数，ESPART 算法的隐私保护性略小于 SMART 算法.

图 4 显示了 ESPART 的隐私保护性的仿真结果，从图中可以看出，$MinDeg$ 越大 ESPART 的隐私保护性越好，然而 $MinDeg$ 越大意味着网络中的数据通信量就越大，在实际应用中，可以采取一种折中的方案，如在保证一定的隐私保护性的同时，尽量减少网络中的数据通信量. 如若希望节点之间的链接被破解的概率低于 2%，在本模拟环境中，$MinDeg$ 取 2 时，节点的隐私数据被暴露的概率就低于 0.02%，可以满足一般的隐私保护要求.

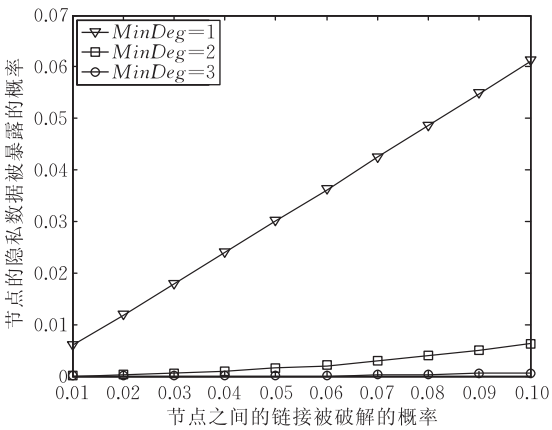


图 4　ESPART 的隐私保护性

将仿真结果 ESPART 与图 1(a)SMART 的隐私保护性[4] 进行对比可以知道，ESPART 中的 $MinDeg$ 对应着 SMART 中的 $J-1$. 在它们数值相等的情况下，ESPART 算法的隐私保护性略低于 SMART 算法，但在 q 值很小的情况下，两种算法的隐私保护性相近，基本可视为同等.

5.2　数据通信量

在 ESPART 算法准备阶段，与 TAG[3] 相同，先从 QS 发出 Hello 信号. 每个节点第一次收到 Hello 信号时，将自己的父节点设置为此 Hello 信号源节点，然后发送一个 Hello 信号，以此建立一棵数据融合树；在向上进行数据融合阶段，每一节点需要在时

间片内向上传递数据.为了简化,在仿真中,我们给每一层分配相同的时间片,并让它们在时间片内随机发送,以避免冲突.ESPART 算法需要每个节点在第一步的融合阶段,记录自己的子节点的个数,并记录于 deg 中,如图 3(a)所示.

由于第一步是准备阶段,其产生的数据通信量与 TAG 相同.在 TAG 中,每个节点都需要发送 2 个数据,即 1 个 Hello 信号和 1 个融合数据.但一旦树型结构固定下来,自己子节点个数记录下来,可以让后面节点间的串通通信阶段减少数据传输量,并且可以重复使用,因此准备阶段产生的数据通信量在 ESPART 算法仿真结果中可以忽略不计.

对于 ESPART 与 SMART 算法,仿真过程中测量的是网络中节点在一次串通期间发送的总数据包个数.对于 TAG 算法,我们衡量的是网络节点在一次建树和融合的过程中发送的总的数据包个数.图 5 说明了 ESPART、SMART 与 TAG 的数据通信量与 $EpochD$ 时间大小无关,ESPART 和 SMART 的数据通信量与 $MinDeg(J)$ 相关,$MinDeg(J)$越大,ESPART 和 SMART的数据通信量越大.

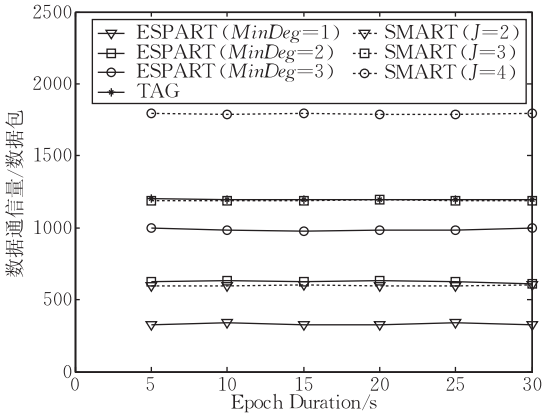


图 5 ESPART 的数据通信量

我们可以从理论上对仿真结果进行分析.对于 ESPART算法,记网络中节点在一次串通通信期间发送的总数据包个数为 $C_E(MinDeg,N)$(N 表示无线传感器网络中节点的数量).同样对于 SMART,记其在一次串通通信期间,网络中节点需要传输的数据包总个数为 $C_S(J,N)$(J 表示 SMART 算法中的分片数目,N 表示无线传感器网络中节点的数量),则有$C_S(J,N)\approx(J-1)\cdot N$.

记 $N_i(i=0,1,\cdots,MinDeg)$表示无线传感器网络数据融合树中,子节点数目为 i 的节点的数量,那么

$$\begin{aligned}C_E(J,N)&<MinDeg\cdot N_0+(MinDeg-1)\cdot N_1+\cdots+\\&\quad 1\cdot N_{MinDeg-1}\\&<MinDeg\cdot(N_0+N_1+\cdots+N_{MinDeg-1})\\&<MinDeg\cdot N.\end{aligned}$$

上式从理论上说明在SMART算法与ESPART算法隐私保护性相近的情况下,从数据通信量(串通花费)方面比较,ESPART 算法要优于 SMART 算法.

从仿真结果可以看出,$C_S(2600)\approx 600$,$C_S(3600)\approx 1200$,$C_S(4600)\approx 1800$,这与理论分析一致.同样可以得到,$C_E(1600)\approx 330$,$C_E(2600)\approx 620$,$C_E(3600)\approx 1000$.因此,在本仿真环境中,ESPART 算法的数据通信量大约是 SMART 算法的 1/2.

5.3 精确度

在无数据丢失的理想环境中,TAG、SMART、ESPART 的数据融合结果都应该能达到 100%的精确度.然而在实际应用中,存在由碰撞、节点间距离过大或时间不足等原因造成数据丢失的现象,从而影响了精确度.我们定义无线传感器网络的精确度为数据融合结果与节点测得真实数据之和的比值.

仿真主要针对 $EpochD$ 的变化对 TAG 和 ESPART算法的影响.在每一个 $EpochD$ 数值下进行 20 次仿真,仿真结果是对它们的均值统计,如图 6 所示.从仿真结果来看,它们的波动较大,各统计数据之间的关系并不能确切表示精确度在一定 $EpochD$ 数值下的相互关系,但是能表示精确度在一定 $EpochD$ 数值下的范围与趋势.

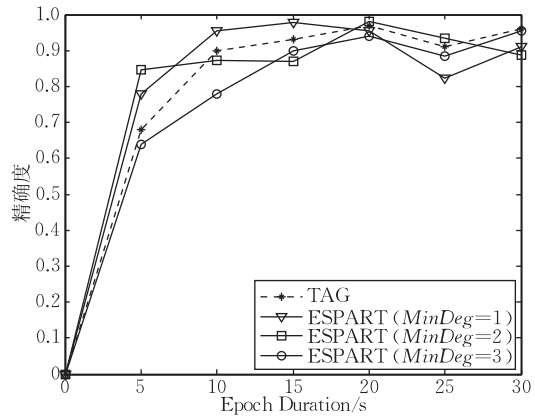


图 6 ESPART 的精确度

从图 6 可以看出,TAG 与 ESPART 算法的精确度都不能达到理想状态下的 100%.并且在 $EpochD$ 值较小的情况下,TAG 与ESPART精确度都随着 $EpochD$ 的增加而增加.当 $EpochD$ 增大到

一定程度时(10s 左右)，精确度都维持在 90%左右. 可以说 TAG 与 ESPART 算法的精确度相对于 $EpochD$ 的变化，其范围与趋势是一致的. 对于 ESPART，$MinDeg$ 对精确度的影响并不大.

对于 SMART 算法[4]，如图 1(c)所示，需要 2 倍于 TAG 的 $EpochD$ 时间才能得到相对精确的结果. 理论上的分析可以得到，由于 ESPART 算法采取了随机分配时间片的方法确定通信时间片，避免了碰撞，且随机数 r 的取值为在(−1，+1)区间内的均匀分布. 由 $Seed_{ij}$ 的计算公式可知，算法可以将串通通信阶段的数据通信量控制在一定范围内. 同时随机数以 1/2 的概率随机选取正负数，以此来降低数据丢失对精确度的影响，得到近似于 TAG 的精确度.

6　结　论

本文提出了一种新的无线传感器网络数据融合隐私保护方案 ESPART，它是对 SMART 方案的一种改进. 相比于 SMART，它可以在有效保护数据隐私的前提下，花费少量的时间与数据通信量，得到精确的数据融合结果.

表 1　SMART 与 ESPART 的性能对比

	SMART	ESPART
隐私保护性	好($J\geq 3$)	好($MinDeg\geq 2$)
数据通信量	大	一般
精确性	好	好
计算量	小	小
鲁棒性	差	好
需要时间	大	小

表 1 总结了 ESPART 相对于 SMART 的优点. ESPART 算法中关于依据数据融合树形结构特性，减少网络数据通信量的思想，可以用在其它无线传感器网络数据融合隐私保护算法中，如 CPDA[4] 与 iPDA[5].

He 提出的 iPDA[5] 与 iCPDA[6] 主要提供了 SMART 与 CPDA 在完整性保护方向上的扩展，分别继承了 SMART 与 CPDA 的缺陷与优势. 同时加入了完整性保护方面的算法，但是它们为了保证数据融合的完整性，都大量增加了网络的数据通信量，并且都需要无线传感器网络中节点足够稠密.

我们今后将展开对 iESPART(an Integrity-Energy-Saving Privacy-preserving AggRegaTion)算法的研究，它是 ESPART 算法在完整性保护方向上的扩展，并着力解决 iPDA 与 iCPDA 遗留的问题.

参　考　文　献

[1] Szewczyk R, Ferencz A. Energy implications of network sensor designs. Berkeley: Berkeley Wireless Research Center Report, 2000

[2] Intanagonwiwat C, Govindan R, Estrin D. Directed diffusion: A scalable and robust communication paradigm for sensor networks//Proceedings of the 6th Annual International Conference on Mobile Computing and Networking. Boston, USA, 2000: 56-67

[3] Madden S, Franklin M J, Hellerstein J M. TAG: A tiny aggregation service for ad-hoc sensor networks//Proceedings of the 5th Symposium on Operating Systems Design and Implementation. New York, USA, 2002: 131-146

[4] He W, Liu X, Nguyen H, Nahrstedt K, Abdelzaher T. PDA: Privacy-preserving data aggregation in wireless sensor networks//Proceedings of the 26th IEEE International Conference on Computer Communications. Anchorage, AK, 2007: 2045-2053

[5] He W, Nguyen H, Liu X, Nahrstedt K, Abdelzaher T. iPDA: An integrity-protecting private data aggregation scheme for wireless sensor networks//Proceedings of the Military Communications Conference. San Diego, CA, 2008: 1-7

[6] He W, Liu X, Nguyen H, Nahrstedt K. A cluster-based protocol to enforce integrity and preserve privacy in data aggregation//Proceedings of the 29th IEEE International Conference on Distributed Computing Systems Workshops. Montreal, QC, 2009: 14-19

[7] Girao J, Westhoff D, Schneider M. CDA: Concealed data aggregation for reverse multicast traffic in Wireless Sensor Networks//Proceedings of the 40th International Conference on Communications. Seoul, Korea, 2005: 3044-3049

[8] Castelluccia C, Mykletun E, Tsudik G. Efficient aggregation of encrypted data in wireless sensor networks//Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services. San Diego, CA, 2005: 109-117

[9] Conti M, Zhang L, Roy S, Pietro R D, Jajodia S, Mancini L V. Privacy-preserving robust data aggregation in wireless sensor networks. Wiley: Security and Communication Networks, 2009, 2: 195-213

[10] Bista R, Jo K J, Chang J W. A new approach to secure aggregation of private data in wireless sensor networks//Proceedings of the 8th IEEE International Conference on Dependable Autonomic and Secure Computing. Chengdu, China, 2009: 394-399

[11] Bista R, Kim H D, Chang J W. A new private data aggregation scheme for wireless sensor networks//Proceedings of the 10th IEEE International Conference on Computer and Information Technology. Bradford, UK, 2010: 273-280

[12] Bista R, Yoo H K, Chang J W. A new sensitive data aggregation scheme for protecting integrity in wireless sensor net-

works//Proceedings of the 10th IEEE International Conference on Computer and Information Technology. Bradford, UK, 2010: 2463-2470

[13] Eschenauer L, Gligor V D. A key-management scheme for distributed sensor networks//Proceedings of the 9th ACM Conference on Computer and Communications Security. Washington, USA, 2002: 41-47

[14] Levis P, Lee N, Welsh M, Culler D. TOSSIM: Accurate and scalable simulation of entire TinyOS applications//Proceedings of the 1st International Conference on Embedded Networked Sensor Systems. Los Angeles, USA, 2003: 126-137

YANG Geng, born in 1961, Ph. D., professor, a senior member of the China Computer Federation. His research interests include computer communication and network, network security, parallel & distributed computing.

WANG An-Qi, born in 1987, M. S. candidate. His research interests include data aggregation and security in wireless sensor networks.

CHEN Zheng-Yu, born in 1978, Ph. D. candidate, lecturer. His research interests include information security and privacy preserving.

XU Jian, born in 1980, Ph. D. candidate, lecturer. His research interests include data aggregation and security.

WANG Hai-Yong, born in 1979, Ph. D. candidate, lecturer. His research interests include data aggregation and security.

Background

Sensing information processing in the Internet of Things (IoT) includes collecting, aggregating, transmitting and analyzing etc. In order to meet the requirements of privacy protection and reliability of decision-making and controlling, a low power consumption privacy-preserving aggregation algorithm is needed to ensure the security in aggregation process and the adaptability of Wireless sensor networks (WSNs). In the recent years, some algorithms about security and privacy-preserving of WSNs have been published, which can be classified into two categories. The first uses homomorphic encryption. It makes nodes aggregate data effectively without using decryption. But it cannot provide any privacy-preserving in terms of QS(Query Server). The second is based on information exchange between nodes to prevent private data exposing, the shortcomings of which are large communication overhead and bad data loss resilience.

This paper focuses on how to design a more effective privacy-preserving aggregation mechanism for WSNs. It proposes an energy-saving privacy-preserving aggregation algorithm: ESPART, which uses characteristic of the data aggregation tree structure to reduce unnecessary communication overhead in SMART, assigns the random time slot to nodes to avoid collision, and limits the scope of collusion data to reinforce data-loss resilience. The theoretical analysis and simulation results show that the ESPART algorithm has better performance in terms of communication overhead and aggregation accuracy than that of the SMART and CPDA.

This research is supported by the National Basic Research Program (973 Program) of China under grant No. 2011CB302903, the National Natural Science Foundation of China under grant No. 60873231 and the Natural Science Foundation of Jiangsu Province, China (grant No. BK2009426).