# 传感器网络中基于路线的隐私保护数据聚集算法

王涛春[1,2],秦小麟[1],张 吉[3],丁有伟[1],陈付龙[2],罗永龙[2]
(1. 南京航空航天大学计算机科学与技术学院,江苏南京 210003;2. 安徽师范大学数学计算机科学学院,安徽芜湖 241002;
3. 南昆士兰大学健康工程与科学学院,澳大利亚图文巴 4350)

**摘 要:** 针对现有隐私保护数据聚集算法依赖某种网络拓扑结构和加解密次数过多的问题,本文提出了一种基于同心圆路线的隐私保护数据聚集算法PCIDA(Privacy-preserving and Concentric-circle Itinerary-based Data Aggregation algorithm).PCIDA沿着设计好的理想路线执行数据聚集,使得算法不依赖网络拓扑结构.PCIDA利用安全通道保证数据的隐私性,避免了数据聚集过程中的加解密运算.PCIDA沿着同心圆并行处理,使得算法数据处理延迟较小.理论分析和实验结果显示,PCIDA在较低通信量和能耗的情况下获得较高的数据隐私性和聚集精确度.

**关键词:** 无线传感器网络;隐私保护;数据聚集;拓扑结构无关;同心圆
**中图分类号:** TP301 **文献标识码:** A **文章编号:** 0372-2112 (2017)06-1334-08
**电子学报URL:** http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2017.06.008

# Privacy-Preserving and Itinerary-Based Data Aggregation Algorithm in Wireless Sensor Networks

WANG Tao-chun[1,2],QIN Xiao-lin[1],ZHANG Ji[3],DING You-wei[1],CHEN Fu-long[2],LUO Yong-long[2]
(1. *College of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics,Nanjing,Jiangsu* 210016,*China*;
2. *College of Mathematics and Computer Science,Anhui Normal University,Wuhu,Anhui* 241003,*China*;
3. *Faculty of Health,Engineering and Sciences,The University of Southern Queensland,Toowoomba,QLD* 4350,*Australia*)

**Abstract:** To solve the problems that the existing privacy-preserving data aggregation relies on a network infrastructure,and data privacy is achieved by excessive encryption process,this paper proposes a privacy-preserving and concentric-circle itinerary-based data aggregation algorithm (PCIDA). Based on a well-designed ideal itinerary for data aggregation, PCIDA is not susceptible to network topology structure. In addition,PCIDA uses secure channel to ensure data privacy with no encryption/decryption operations during data aggregation. PCIDA performs data aggregation in parallel along with well-designed concentric-circle itineraries to achieve small delivery delay. Theoretical analysis and experimental results show that PCIDA enjoys low communication overhead and energy consumption,yet high safety and accuracy.

**Key words:** wireless sensor networks;privacy-preserving;data aggregation;infrastructure-free;concentric-circle

## 1 引言

无线传感器网络(Wireless Sensor Networks,WSNs)是由大量部署在物理世界具有感知、数据处理和通信功能的传感器节点构成.节点一般通过电池供电,不可更换,所以无线传感器网络关注的首要问题是节省节点能耗以延长网络使用寿命,数据聚集[1~3]是减少数据冗余降低通信量的重要技术.同时WSNs面临严重的数据泄露问题,如何保证感知数据的隐私性是拓展无线传感器网络应用领域的关键因素,是无线传感网络研究中的一个热点问题.

通常隐私保护数据聚集算法利用加密技术保证感知数据的隐私性,但节点需要进行大量的加解密运算,从而造成严重的处理延迟和能量能耗.文献[4]提出了

收稿日期:2016-08-16;修回日期:2016-09-20;责任编辑:覃怀银
基金项目:国家自然科学基金(No.61402014,No.61373015,No.61672039,No.61572036);国家教育部高等学校博士学科点专项科研基金资助项目(No.20103218110017);中央高校基本科研业务费专项基金项目(No.NP2013307);安徽省自然科学基金项目(No.1508085QF133)

一种无需加解密的安全加聚集函数,Sink 节点与每个节点共享一个随机数,节点通过加上随机数隐藏感知数据真实值,Sink 节点将最终的聚集结果减去所有节点随机数得到真实的聚集结果. 由于 Sink 节点很难追踪哪些节点因为信道冲突没有完成数据传输,使得聚集结果可能减去额外随机数而造成严重的偏差. 现有算法一般基于某种网络拓扑结构(树或簇),但无线传感器网络的动态性特点,使得网络拓扑结构维护代价较大. 文献[2]提出了一种结构无关的聚集算法 IWQE,算法沿着一条或多条动态生成的路线收集节点的数据,IWQE 能够有效地减少了网络结构维护能耗,但该算法没有考虑数据的隐私性.

针对上述问题,本文提出基于同心圆路线的隐私保护数据聚集算法 PCIDA. 该算法利用安全通道来防窃听攻击且避免了加解密运算;通过实时确定聚集节点,使得 PCIDA 不依赖网络拓扑结构,并利用同心圆路线并行处理降低处理延迟;算法通过对数据节点数据进行切分操作[4]使得聚集节点或其它节点不能获取数据节点的数据,且节点之间不需要提前部署共有信息,因而网络具有良好的扩展性。

## 2　相关工作

数据聚集是降低通信量的主要技术,隐私保护是 WSNs 应用于某些领域的关键因素,提出了很多高效的隐私保护数据聚集算法[5~10]. He 等人[5]提出了安全的加数据聚集算法 SMART 和 CPDA,SMART 利用切分重组技术实现隐私保护数据聚集,CPDA 利用随机数和多项式的代数性质实现隐私保护的数据聚集算法. Ozdemir 等人[6]提出了隐私保护数据聚集算法 PRDA,PRDA 算法中,传感器节点将数据表示成多项式函数以减少数据通信量. Lin 等[7]提出了一种多个应用场景的隐私保护数据聚集方案 CDAMA. Chen 等[9]提出了可恢复的隐私数据聚集算法 RCDA. Rezvani 等人[10]提出了一种安全的数据聚集算法,算法改进了迭代过滤并降低了串谋对数据的影响.

由于无线传感器网络结构具有动态性,使得基于树或簇的聚集算法需要实时维护网络的拓扑结构,Xu 等人[2]提出了一种基于路线的数据聚集算法 IWQE,该算法主要由三个步骤组成:Sink 发送聚集请求、聚集区域进行聚集处理、节点将聚集结果返回给 Sink 节点,但 IWQE 算法没有考虑感知数据的隐私性和链路质量等问题. Han 等[11]提出了基于路线的 KNN 查询处理算法.

## 3　PCIDA 算法

隐私保护的数据聚集算法 PCIDA 包括一个 Sink 节点和具有 $N$ 个节点的聚集区域(二维圆形窗口),在保证感知数据隐私的情况下,Sink 节点获得聚集区域内所有节点的聚集结果. 所有节点维护其自身和安全链接邻居节点(Secure Link Neighbors,SLN)的位置信息,其中 SLN 节点指的是两邻居节点至少共享一个密钥. 所有节点通过周期性交换信息来维护 SLN 节点列表. 本文使用的主要符号及描述见表 1.

**表 1　符号及其描述**

| 符号 | 描述 |
|---|---|
| $R$ | 传感器节点的传输半径 |
| Area | 聚集区域面积 |
| $N$ | 聚集区域内的节点数 |
| $n_s$ | 传感器节点具有安全链接邻居节点平均个数 |
| $e_T$ | 传感器节点发送 1bit 数据所需能耗 |
| $e_R$ | 传感器节点接收 1bit 数据所需能耗 |
| $E_a$ | 通信所需总能耗 |
| $E_t$ | 数据聚集算法总能耗 |
| $D$ | 聚集节点发布聚集请求和接收聚集数据的时间 |
| $f_{i,j}$ | 传感器节点 $n_i$ 和 $n_j$ 之间的距离 |
| $R_d$ | 感知数据范围为$[0,R_d]$ |
| $W$ | 路线宽度 |

### 3.1　同心圆线路设计

为了保证能够聚集区域内所有节点的感知数据,路线宽度满足 $W \leqslant \sqrt{3}R/2$[2]. 在此基础上设置理想路线,从第一个聚集节点开始,沿着理想路线确定下一个聚集节点,如此继续直到最后一个聚集节点,形成聚集节点序列,即实际路线. 由于实际路线是沿着理想路线实时确定的,不依赖于网络结构. 如图 1 所示,3 条蓝色虚线将聚集区域分成 4 个聚集子区域,灰色虚线表示理想路线,沿着理想路线的聚集节点(黑色圆点)形成实际路线,灰色圆点表示数据节点,每个聚集节点只聚集本子聚集区域的数据节点. 聚集节点发布聚集请求和接收数据所需时间基本固定,因此算法总的延迟由理想线路总长度决定,具体为式(1),其中 $d_A$ 为聚集节点之间的平均距离.

$$L = \frac{\sum_{i=1}^{\left|\frac{\sqrt{\text{Area}/\pi}}{W}\right|} 2\pi \times W \times i}{d_A} \times D \tag{1}$$

虽然通过平行发布能够有效的减少聚集延迟,但也可能引发信道冲突. 因此,设计并行路线必须考虑信道发生冲突的概率,且要保证多路线数据聚集结果能够有效的返回给 Sink 节点. 因此,本文提出一种基于同心圆路线并行处理数据聚集. 每环同心圆起始聚集节点不仅沿着理想路线(逆时针方向)确定下一个聚集节点,同时确定外一层环起始聚集节点,为了降低由于并行操作造成的信道冲突,选择外一层环的起始聚集节点时尽量选择与其理想路线反向(顺时针)距离尽量远,

尽量贴近理想路线的节点. 如图 2 所示,第二层环起始聚集节点 $A_{21}$ 沿理想路线选择 $A_{22}$ 作为下一个聚集节点,且选择第三层环的 $A_{31}$ 作为该环起始聚集节点,使得每

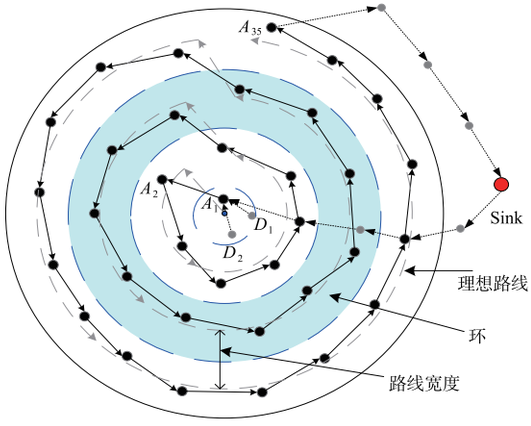



图1 基于路线的数据聚集

$$L_{\mathrm{CCI}}=\left(\frac{2\pi\times W\times\left|\frac{\sqrt{\mathrm{Area}/\pi}}{W}\right|}{d_A}+\left|\frac{\sqrt{\mathrm{Area}/\pi}}{W}\right|\right)\times D \tag{2}$$

### 3.2 算法思想

由于基于网络拓扑结构的数据聚集算法容易受到网络结构动态变化的影响,且网络拓扑结构需要大量的维护能耗. 同时为了在数据聚集过程中保证数据的隐私性,本文提出了聚集算法 PCIDA. 算法主要分为 5 个阶段:(1) 路线设计阶段:根据聚集区域大小、通信半径等参数确定网络理想线路. 并利用 3.1 节设计的路线并行处理;(2) 初始化阶段:对已预设置密钥的邻居节点之间建立安全通道;(3) 聚集请求阶段:利用位置路由协议将 Sink 节点的聚集请求传输给起始聚集节点;(4) 数据聚集阶段:聚集区域内所有数据节点数据;(5) 聚集结果返回阶段:对最终聚集结果加密并返回给 Sink 节点.

#### 3.2.1 初始化阶段

节点从具有 $K$ 个密钥的大密钥池中随机选取 $k$ 个密钥[12],并与安全链接邻居节点建立安全通道. 例如,节点 $n_i$ 与其 SLN 节点 $n_j$(具有相同密钥 $k_{i,j}$)建立安全通道,具体为:节点 $n_i$ 利用密钥 $k_{i,j}$ 对产生的随机数 $d_{i,j}$ 加密,并将密文传输给节点 $n_j$,节点 $n_j$ 利用密钥 $k_{i,j}$ 解密得到随机数 $d_{i,j}$(随机数 $d_{i,j}$ 即为安全通道,且 $d_{i,j}=d_{j,i}$). 节点与其所有 SLN 节点建立安全通道,且节点需要维护其 SLN 节点列表.

#### 3.2.2 数据聚集阶段

数据节点感知数据并将数据发送给聚集节点,聚集节点接收数据并进行聚集操作,最后一个聚集节点环可并行处理. 从图 2 可知,该并行处理总的延迟由最内层环到最外层环路线和最外层环理想路线长度决定,该路线的聚集延迟见式(2).

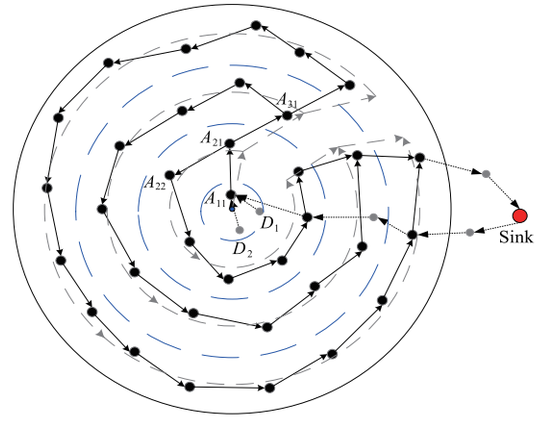



图2 同心圆并行路线

得到最终的聚集结果. 该阶段具体操作如下:

(1) 聚集节点选择

PCIDA 采用同心圆路线并行处理,每环内的聚集节点分成三类:首个聚集节点为起始聚集节点、最后一个聚集节点为终止聚集节点、其余聚集节点为中间聚集节点,不同聚集节点具有不同的操作.

($a$) 起始聚集节点需要在环内选择下一个聚集节点、选择外一层环起始聚集节点和向数据节点广播聚集请求。具体为:起始聚集节点在 SLN 节点列表中选择与理想线路尽可能贴近、前进距离尽可能远(逆时针)的节点作为下一个聚集节点,具体选择式为 $S_i^A(j)=\gamma\times f_{i,j}^d+\delta\times f_j^l$,其中 $\gamma$ 和 $\delta$ 分别为向前距离 $f_{i,j}^d$ 和贴近理想路线 $f_j^l$ 的权重系数($\gamma+\delta=1$);起始聚集节点选择外一层环与理想路线尽可能贴近、后退距离尽可能远(顺时针)的节点作为外一层环起始聚集节点,选择式为 $S_i^{\mathrm{nextc}}(j)=\gamma_1\times f_{i,j}^d+\delta_1\times f_j^l$;聚集节点向邻居节点广播聚集请求信息,接收并聚集感知数据,并将当前聚集结果等信息传输给下一聚集节点.

($b$) 中间聚集节点:选择下一个聚集节点的标准与起始聚集节点相同为式 $S_i^A(j)$,给邻居节点广播聚集请求信息和给下一个聚集节点传输聚集结果等信息.

($c$) 终止聚集节点:给邻居节点广播聚集请求信息,并将聚集结果传输给外一层环终止聚集节点.

(2) 数据切分与重组

当收到聚集节点的聚集请求信息后,符合响应请求的节点,称之为数据节点. 为了保证感知数据的隐私性,数据节点 $D_i$ 首先对感知数据 $v_i$ 随机分成 $J_i$ 片,并将 $J_i-1$ 片数据通过安全通道传输给 SLN 节点,再对其它数据节点发送过来的分片数据和自身保留的分片数

据进行重组操作,最后将重组结果发送给目的聚集节点.

由于在破解链路安全概率确定的情况下,感知数据的隐私性由数据节点的切片数和收到其它数据节点切片数确定,因此,PCIDA 中数据节点根据已收到的切片数动态确定自身数据切片个数,使得所有数据节点感知数据的切片数与接收到的切片数之和尽可能相等,优点是总切片数(通信量)一定的情况下,感知数据能够获得尽可能高的平均隐私性.

(3) 数据聚集

当数据节点接收到其它数据节点发送的切片数据后,对所有切片数据进行重组操作,再将重组的结果传输给目的聚集节点. 聚集节点接收到数据节点重组数据和上一个聚集节点发送的部分聚集结果后,对所有数据(包括自身感知数据)进行聚集操作,再将聚集结果发送给下一个聚集节点,如此继续,直到遍历完该同心圆区域内的所有节点,计算出该同心圆的聚集结果并传输给外一环的终止聚集节点.

### 3.3 性能分析

#### 3.3.1 安全性分析

在分析安全通道安全性的基础上分别对数据节点和聚集节点感知数据的隐私性进行分析.

(1)安全通道:其它节点或攻击者可以通过解密或猜测的方式来获取随机数 $d_{i,j}$:($a$) 对于其它邻居节点来说,通过拥有的密钥解密获得随机数的概率为 $p = k/K$[12];(2) 随机数 $d_{i,j}$ 在范围[0,$r$]中是均匀分布的,其中 $r = R_d * N$,所以正确猜测的概率是 $1/r$. 因此,安全通道被泄露的概率为

$$P_r = \min((\sigma \times p \times (n_s - 1)), r) \tag{3}$$

其中 $n_s$ 为 SLN 节点数,参数 $\sigma$ 为安全通道重建频率对安全性影响的系数.

(2) 数据节点:要得到数据节点 $D_i$ 感知数据 $v_i$ 值,攻击者必须得到 $D_i$ 传输的 $J_i - 1$(出度)片数据和重组数据值、其它数据节点传输给 $D_i$ 的 $J_i^r$(入度)片数据,因此,$D_i$ 感知数据被泄漏的概率为

$$P_i^D = P_r^{J_i - 1} \times P_r^{J_i^r + 1} = P_r^{J_i + J_i^r} \tag{4}$$

(3) 聚集节点:要推导出聚集节点 $A_i$ 的感知数据 $v_i$ 值,攻击者必须得到所有传输给 $A_i$ 的重组值、可能有的分片数据和上一个聚集节点的部分聚集值,以及 $A_i$ 的部分聚集结果发送给下一个聚集节点,所以感知数据 $v_i$ 泄漏的概率为式(5),其中 $n_i^D$ 表示向 $A_i$ 传输数据的数据节点数.

$$P_i^A = P_r^{n_i^D} \times P^{J_{A_i} + 1} = P^{n_i^D + J_{A_i} + 1} \tag{5}$$

在切片总数确定的情况下,当所有数据节点的出入度之和相等时,数据节点感知数据平均泄漏的概率最低. 虽然在实际执行过程中很难保证所有数据节点出入度之和相等,但当所有数据节点的出入度之和与平均值越接近,则其感知数据平均泄漏的概率越低. 因此,数据节点通过动态确定切片数 $J$,使得出入度总数尽量接近平均值,在不增加通信量的情况下,能够最大程度提高感知数据的平均隐私性.

#### 3.3.2 通信量

PCIDA 通信量主要包括建立安全通道和执行数据聚集两部分组成. 每条安全通道建立需要传输密文以及返回确认信息,密文数据的位长为 $\bar{e}$. 节点 $n_i$ 需要和 $n_i^s$ 个节点建立安全通道,且安全通道是双向的,所以安全通道总的通信量为

$$T_s = \frac{1}{2}\sum_{i=1}^{N}(\bar{e} + 1) \times n_i^s \tag{6}$$

PCIDA 中感知数据、切片数据和聚集结果的数据位长相等,为 $\bar{d} = |\log(R_d \times N)|$. 聚集节点广播聚集请求、发送部分聚集结果和接收重组数据和可能有的切片数据,聚集请求位长为 $\bar{q}_a$. 聚集节点个数为 $A_{\text{num}}$,聚集节点 $A_i$ 接收到的重组数个数为 $n_i^D$ 和切片数为 $J_{A_i}^r$,其通信量为

$$T_A = \sum_{i=1}^{A_{\text{num}}}(\bar{q}_a + (n_i^{D_{\text{num}}} + 1 + J_{A_i}^r) \times \bar{d}) \tag{7}$$

数据节点 $D_i$ 需发送 $J_i - 1$ 片数据和一个重组结果,接收 $J_i^r$ 片数据,数据节点个数为 $D_{\text{num}}$,因此,其通信量为

$$T_D = \sum_{i=1}^{D_{\text{num}}}((J_i + J_i^r) \times \bar{d} + \bar{q}_a) \tag{8}$$

聚集区域内节点总数为 $N(N = D_{\text{num}} + A_{\text{num}})$,每个数据节点接收聚集请求消息. 由于发送的切片总数与重组数据个数之和与总切片数量相等,因此,总通信量可简化为

$$T_A = N \times \overline{q_a} + \bar{d} \times (2 \times \sum_{i=1}^{D_{\text{num}}} J_i + A_{\text{num}}) \tag{9}$$

由式(9)可知,通信量与切片数据总个数有关,以及与数据节点和聚集节点所占比例有关. 为了进一步提高数据的隐私性,每隔一段时间需要重建安全通道,假设每段时间进行 $\beta$ 轮数据聚集,则每轮数据聚集的平均通信量为

$$T = \frac{1}{\beta}(T_s + \beta \times T_A) \tag{10}$$

#### 3.3.3 能耗

传感器节点能耗主要包括通信能耗和计算能耗,相比加解密计算能耗,加模计算能耗可忽略不计. 由于 PCIDA 只在安全通道建立阶段执行加解密运算,且每条安全通道需执行一次加密和一次解密计算,因此,加解密次数具体为

$$N_{\text{dec}} = N_{\text{enc}} = \frac{1}{2}\sum_{i=1}^{N} n_i \tag{11}$$

根据上述通信量的分析,可得出数据节点和聚集

节点通信量能耗,具体为

$$E_a = \sum_{i=1}^{D_{num}} (J_i \times \bar{d} \times e_T + (J_i^r \times \bar{d} + \overline{q_a}) \times e_R) + \sum_{i=1}^{A_{num}} (\overline{q_a} \times e_T + (n_i^{D_{num}} + 1 + J_{A_i}^r) \times \bar{d} \times e_R) \quad (12)$$

因此,每轮数据聚集的平均能耗为

$$E_t = \frac{1}{\beta}((E_{dec} + E_{nec}) \times N_{dec} + \beta \times E_a) \quad (13)$$

## 4 仿真实验

为了对算法的通信量、能耗、精确度和隐私性等性能进行比较,本节仿真实现 PCIDA、SMART 和 PRDA. 实验环境为 Core i3 CPU、4GB 内存、Win7 操作系统、VS. NET 2010 和 Matlab. 通信能耗分别为 $e_T = 0.72\mu J$ 和 $e_R = 0.81\mu J$,RC4 对 10bit 数据加解密能耗为 $8.92\mu J$[13]. 实验结果是执行 20 轮的平均值,每轮随机产生一个无线传感器网络拓扑. 其它参数默认值如表 2 所示.

**表 2 实验参数默认值**

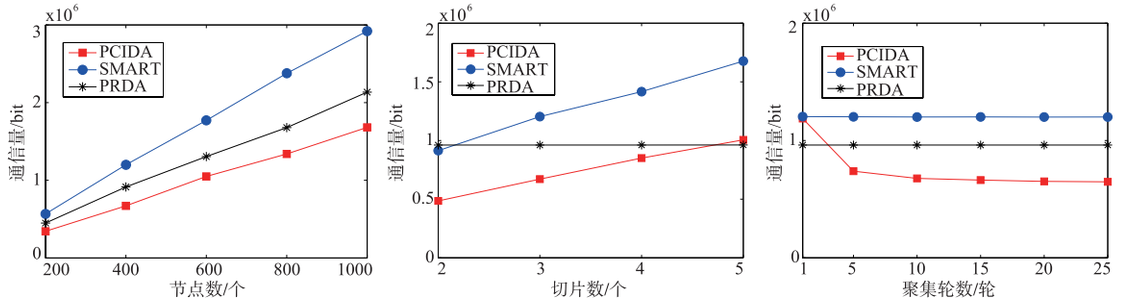
| 参数名 | 参数值 |
|---|---|
| 网络覆盖区域 | 100m × 100m |
| 网络节点分布 | 随机均匀分布 |
| 节点通信半径 | 10m |
| 节点数 | 400 |
| 感知数据消息大小 | 40byte |
| 聚集请求消息大小 | 20byte |
| 聚集区域占网络覆盖区域的百分比 | 100% |
| 网络宽度占通信半径的比率 | $\sqrt{3}/2$ |

### 4.1 通信量及能耗

实验 PCIDA、SMART 和 PRDA 在不同的节点数、不同的平均切片数和每周期不同数据聚集轮数情况下,算法总的通信量和能耗.

(1) 通信量

图 3(*a*)显示不同节点数目下,3 种算法所需要的通信量. PRDA 中节点不需要进行切片操作,且用多项式系数来表示数据以进一步减少数据传输量,所以通信量比 SMART 低. 但 PRDA 中不参加数据聚集的节点需要通过多跳传输密文给 Sink 节点,同时聚集节点需要多跳传输多项式系数,因而 PCIDA 通信量最低,且随着节点数 $N$ 的增大,通信量减少的幅度更大. 图 3(*b*)显示切片数 $J$ 对 3 种算法所需要的通信量,由该图可知,PCIDA 的通信量比 SMART 低 35% 以上,且 $J$ 的值越大,PCIDA 通信量比 SMART 减少得越多,主要原因是 SMART 需要对所有节点进行切片,而 PCIDA 只需要对数据节点进行切片. PRDA 不需要进行切片操作,所以通信量没有变化,总体上,PCIDA 通信量优于 PRDA,只有在切片数较大时,PRDA 通信量略低于 PCIDA. 从图 3(*c*)可知,只有当每轮聚集都需重新建立安全通道($\beta = 1$)时,PCIDA 的通信量比 SMART 和 PRDA 稍高. 随着 $\beta$ 值越大,PCIDA 平均每轮数据聚集需要的通信量越低. 当 $\beta$ 值较大时,PCIDA 需要通信量减少非常明显.

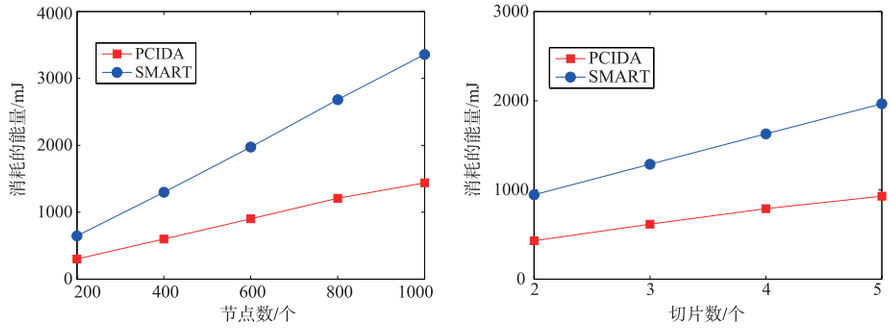



(*a*) 节点数*N*的影响 (*b*) 切片数*J*的影响 (*b*) 聚集轮数$\beta$的影响

图3 通信量

(2) 能量消耗

图 4(*a*)显示节点数对 PCIDA 和 SMART 能耗的影响,由图可知,相比较通信量,PCIDA 能耗比 SMART 降低幅度更大,达到 50%,而且随着 $N$ 的增大,降低幅度也随之增大,主要原因是 PCIDA 只在安全通道建立阶段需要加解密操作. 因此,即使不考虑网络拓扑结构维护开销的情况下,PCIDA 花费更少的能耗来保证感知数据的隐私性. SMART 随着切片数的增加,加解密次数增加,而 PCIDA 的加解密次数与切片数无关,因此,随着切片数 $J$ 的增大 PCIDA 比 SMART 节省更多能耗,如图 4(*b*)所示.




（a）节点数$N$的影响　　（b）切片数$J$的影响

图4　能耗

## 4.2　隐私保护

图 5 显示 PCIDA、SMART 和 PCIDA_Random 感知数据泄漏的概率,其中 PCIDA_Random 表示没有应用切片数优化确定机制,即数据节点切片数完全随机,不考虑节点出入度之和的均衡性. 实验结果显示,在破解链路安全的概率 $p_l$ 较低时 PCIDA 和 SMART 感知数据泄漏的概率近似,都具有较高的隐私性. 但随着 $p_l$ 的增大,PCIDA 具有更好的隐私性,特别 $p_l$ 大于等于 0.8 时,PCIDA 数据的隐私性比 SMART 具有明显的优势. 同时,由于 PCIDA_Random 没有综合考虑节点的出度和入度,使得节点间隐私度相差较大,造成数据平均隐私性较低,所以 PCIDA_Random 数据泄漏率最高.

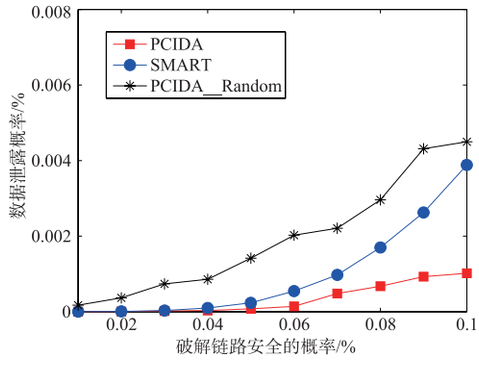



图5　私有性比较图

## 4.3　聚集精确度

由于网络无线通道的冲突性、数据处理的延迟性导致传输信息的丢失和节点的失效性,使得聚集结果的精确性受到影响. 聚集精确度为算法聚集结果与实际的感知数据和之间的比率. 图 6 显示了 PCIDA 和 SMART 在不同周期长度得到的聚集结果精确度. 从图中可知,时间周期越长,聚集精确度越高. 此外,PCIDA 聚集精确度比 SMART 更高,由于 SMART 需要对每片数据进行加解密操作,而 PCIDA 不需要对分片数据进行加解密操作,减少了处理延迟,所以当时间周期相对较小时,PCIDA 具有更明显的优势.

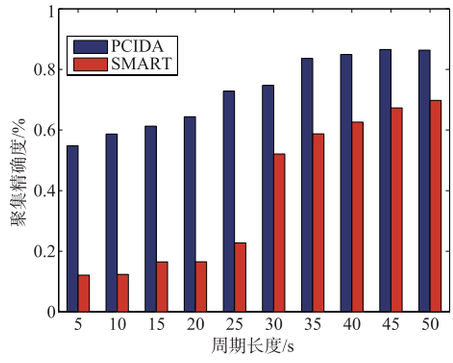



图6　周期长度对数据聚集精确度的影响

## 4.4　时间延迟

图 7 显示了 PCIDA 和 PCIDA_SI(采用串行路线,如图 1 所示)的数据聚集延迟. 可以看出 PCIDA_SI 的延迟值比较大,主要原因是 PCIDA 采用基于同心圆并行路线,极大地缩短了线路的总长度.

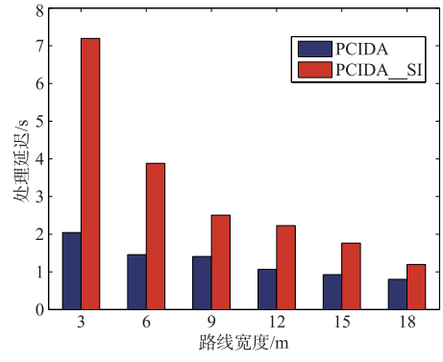



图7　路线宽度对处理延迟的影响

## 5 结论

无线传感器网络中,具有隐私保护的高效的数据聚集算法是一个挑战性的问题. 本文提出了隐私保护的数据聚集算法 PCIDA,该算法采用基于路线的聚集请求和数据聚集,使得算法的执行与网络的拓扑结构无关,不仅节省了网络拓扑结构维护能耗,且比较适用于拓扑结构频繁变化的无线传感器网络. 同时,所有数据通过安全通道进行传输,在不需要执行加解密操作的情况下保证了数据的隐私性. 由于在聚集过程中无任何加解密操作,不仅节省了计算能耗,且缩短了数据处理所需时间,提高了聚集精确度. 实验结果显示 PCIDA 在增加少量通信量和能耗的情况下,保证了数据的隐私性,与现有安全数据聚集算法相比,PCIDA 需要更少的通信量和能耗获得较高的聚集精确度,且不受网络结构影响.

### 参考文献


[1] Bagaa M, Younis M, Djenouri D, et al. Distributed low-latency data aggregation scheduling in wireless sensor networks[J]. ACM Transactions on Sensor Networks, 2015, 11(3):49.

[2] Xu Y, Lee W, Xu J, et al. Processing window queries in wireless sensor networks[A]. The 22nd International Conference on Data Engineering[C]. Atlanta, USA: IEEE, 2006. 70 – 80.

[3] 徐佳,冯鑫,杨富贵,等. 最大化最小能耗概率的移动 Sink 无线传感器网络数据收集方法[J]. 电子学报, 2015,43(12):2470 – 2475.
Xu Jia, Feng Xin, Yang Fu-gui, et al. A data collection method by maximizing minimum probability of energy consumption for mobile Sink based WSNs[J]. Acta Electronica Sinica, 2015, 43(12):2470 – 2475. (in Chinese)

[4] 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, USA: IEEE, 2005. 109 – 117.

[5] He W, Liu X, Nguyen H, Nahrstedt K, Abdelzaher T. PDA: Privacy-preserving data aggregation for information collection[J]. ACM Transactions on Sensor Networks, 2011, 8(1): Article No. 6.

[6] Ozdemir S, Miao P, Yang X. PRDA: polynomial regression-based privacy-preserving data aggregation for wireless sensor networks[J]. Wireless Communications and Mobile Computing, 2015, 15(4):615 – 628.

[7] Lin Y H, Chang S Y, Sun H M. CDAMA: concealed data aggregation scheme for multiple applications in wireless sensor networks[J]. IEEE Transactions on Knowledge and Data Engineering, 2013, 25(7):1471 – 1483.

[8] Parmar K, Jinwala D C. Concealed data aggregation in wireless sensor networks: A comprehensive survey[J]. Computer Networks, 2016, 103:207 – 227.

[9] Chen C M, Lin Y H, Lin Y C, et al. 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.

[10] Rezvani M, Ignjatovic A, Bertino E, Jha S. Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks[J]. IEEE Transactions on Dependable and Secure Computing, 2015, 12(1): 98 – 110.

[11] Han Y, Tang J, Zhou Z B, et al. Novel itinerary-based KNN query algorithm leveraging grid division routing in wireless sensor networks of skewness distribution[J]. Personal and Ubiquitous Computing, 2014, 18(8):1989 – 2001.

[12] Eschenauer L, Gligor V D. A key-management scheme for distributed sensor networks[A]. The 9th ACM Conference on Computer and Communications Security[C]. Washington, USA: ACM, 2002. 41 – 47.

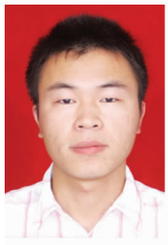
[13] Groat M M, He W, Forrest S. KIPDA: k-indistinguishable privacy-preserving data aggregation in wireless sensor networks[A]. The 30th IEEE International Conference on Computer Communications[C]. Shanghai, China: IEEE, 2011. 2024 – 2032.


### 作者简介


**王涛春** 男,1979 年生于安徽省芜湖市. 现为安徽师范大学数学计算机科学学院副教授、硕士生导师. 主要研究方向为隐私保护、无线传感器网络和群智感知.
E-mail: wangtc@nuaa.edu.cn

**秦小麟** 男,1953 年生于江苏省南京市. 现为南京航空航天大学计算机科学与技术学院教授、博士生导师. 主要研究方向为分布式数据管理与安全,时空数据管理.
E-mail: qinxcs@nuaa.edu.cn

**张　吉**　男,1977 年生于四川省成都市. 现为澳大利亚南昆士兰大学计算机科学教授、博士生导师、澳大利亚奋进学者、澳大利亚昆士兰学者. 主要研究方向为大数据分析、数据挖掘,、信息隐私保护及安全.
Email:Ji. Zhang@ usq. edu. au



**丁有伟**　男,1987 生于江苏省宿迁市,现为南京航空航天大学计算机科学与技术学院博士研究生. 主要研究方向为能量高效数据管理、云计算.
E-mail:dingyouwei@ nuaa. edu. cn

**陈付龙**　男,1978 年生于安徽省霍邱县,现为安徽师范大学数学计算机科学学院教授、硕士生导师. 主要研究方向为嵌入式计算和普适计算、高性能计算机体系结构.
E-mail:Long005@ mail. ahnu. edu. cn

**罗永龙**　男,1972 年生于安徽省太湖县,现为安徽师范大学数学计算机科学学院教授、博士生导师、安徽省学术和技术带头人. 主要研究方向为信息安全、空间数据处理.
E-mail:ylluo@ ustc. edu. cn