数据融合中支持隐私保护的完整性动态验证算法

陈 伟 杨 龙 于 乐

(南京邮电大学计算机学院 南京 210023) (江苏省无线传感网高技术研究重点实验室 南京 210003)

摘 要 隐私暴露、信息篡改、虚假数据注入都是无线传感器网络数据融合中面临的严峻挑战,在保护数据隐私性的同时进行完整性验证是数据融合技术研究的热点之一。提出了一种新的支持隐私保护的动态完整性验证算法 PDI (Privacy-preserving Dynamic Integrity-verification algorithm),它可以在实现数据隐私保护的基础上侦测到信息被非法篡改。PDI 算法使用数据扰动进行数据隐私保护,同时根据现有网络结构动态生成监测节点进行数据的完整性验证,融合过程中篡改过的虚假数据能够更快地被检测并丢弃。仿真实验结果显示,PDI 算法可使用较少的通信量和计算量实现隐私保护和完整性验证。

关键词 无线传感器网络,数据融合,隐私保护,完整性验证,数据扰动

中图法分类号 TP393-08 **文献标识码** A

Privacy-preserving Dynamic Integrity-verification Algorithm in Data Aggregation

CHEN Wei YANG Long YU Le

(School of Computer Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

(Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210003, China)

Abstract Privacy exposure, information tampering and false data injection are serious challenges in wireless sensor network data aggregation. How to protect the privacy and integrity of data has become a hot research issue in data aggregation. To detect any malicious data tampering and protect data privacy in data aggregation, a novel Privacy-preserving dynamic Integrity-verification(PDI) algorithm was proposed. The PDI algorithm uses the data is perturbation to protect data privacy. In order to protect the integrity of privacy data, the PDI algorithm generates monitoring nodes between the current aggregator and its parent node depending on the network structure dynamically. Therefore, if the data is tampered by the aggregator, it can be detected at the early stage. The simulation results show that the PDI algorithm can realize privacy preserving and integrity verification with less communication and computation overhead.

Keywords Wireless sensor network, Data aggregation, Privacy preserving, Integrity verification, Data perturbation

1 引言

无线传感器网络由大量的传感器节点组成,每个节点的能量、计算、存储、感应及通信能力有限。为了节省能量和通信带宽,减少原始数据的发送,传感器节点需要在网内协同处理所收集的原始数据,数据融合就是在网内处理的方法之一。

由于无线传感器网络的开放特性,数据融合过程中数据容易被捕获和窃听,如果攻击者破解了无线链路或捕获了网内节点,网络中的数据便暴露了。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,通常使用加密或数据扰动等方法对原始数据进行保护。

但是,攻击者通过俘获或复制传感器节点等手段成为网络的参与者来实现内部攻击,对数据进行恶意插入、修改或者删除。由于数据已进行隐私保护,内部合法节点在接收到这样的数据时无法判断完整性是否被破坏,因此需在进行数据隐私保护的同时,对接收的数据进行完整性验证,实现在传感数据的真实内容不可知的情况下检测到非法篡改。

本文提出了一种支持隐私保护的完整性动态验证算法 PDI,算法采用数据扰动的方式对数据进行隐私保护,通过引入多个监测节点对融合节点进行监督,防止节点对数据进行非法篡改,有效地验证了完整性。该算法根据现有的树形结构自适应地形成监测节点对,具有很好的灵活性和通用性;同时,PDI 算法采用数据干扰的方式进行隐私保护,所以不需要对传输数据进行逐跳加解密运算,可以有效地减少通信和计算代价;此外,PDI 算法数据在传输过程中会逐跳进行完整性验证,所以虚假数据能够更快地被验证丢弃。

本文第 2 节讨论相关工作,介绍无线传感器网络中主流的数据隐私保护以及完整性验证算法;第 3 节描述本文所采用的系统模型;第 4 节详细介绍支持隐私保护的完整性动态验证算法 PDI;第 5 节通过理论方法和仿真实验对 PDI 算法进行性能分析,并与同类算法进行对比。

2 相关工作

在无线传感器网络中,数据融合能够有效去除冗余信息、

到稿日期:2012-09-02 返修日期:2012-12-23 本文受国家自然科学基金(61202353,61272084)资助。

陈 伟(1979—),男,博士,副教授,CCF 会员,主要研究方向为网络安全,E-mail:chenwei@njupt.edu.cn;杨 龙(1987—),男,硕士生,主要研究方向为无线传感器网络安全;于 乐(1990—),男,硕士生,主要研究方向为数据隐私保护。

减少传输量，从而极大地节省了传感器节点能量，延长了网络生命周期。文献[1]中提出的 TAG 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。同时，由于传感器节点通常暴露在野外，容易遭受各种各样的攻击，因此，如何在保证传感器节点的能源高效性的同时确保融合数据传输的安全性，成为无线传感器网络中颇受关注的研究领域。近年来，国内外的一些研究者提出了各种不同的方案和协议，以保证数据融合过程中的安全性。

针对 WSNs 数据隐私保护的问题，He 等提出了算法 PDA[2]，其中包含了两种数据隐私保护方法：SMART 与 CPDA，在 SMART 方法中，源节点将数据切割成片后加密发送，因而中继节点无法获得完整的数据，从而实现了对数据隐私的保护。SMART 与 CPDA 均只支持 SUM 聚集操作，同时由于使用逐跳加密机制，计算与通信代价较高。此后，He 等人提出了 iPDA[3]算法，算法的隐私性类似于 SMART 通过切片的思想实现；然后通过两棵节点无交集的聚焦树，对比冗余数据来验证数据的完整性，冗余数据也带来了额外的通信开销。Feng 等也通过引入干扰数的方式实现数据隐私保护[4]，并进一步提出了一系列改进算法来提高通信性能。

Y. Yang，X. Wang 等提出了 SDAP(Secure Hop-by-Hop Data Aggregation protocol)[7]。SDAP 采用随机方法把拓扑树动态地分成多个同样大小的局部子树，尽可能地减少一个树上的上层节点的数量，从而消除由上层节点引起的安全威胁。在通常的逐跳融合方法基础上添加一个承诺性能，以帮助基站验证融合数据的正确性。Przydatek 等也设计了 SIA[6]安全信息融合框架，SIA 在融合结果已被篡改的情况下，用户应该以较高概率拒绝不正确的融合结果。

为了同时实现数据融合过程中的数据隐私性、完整性，Chuang Wang 等提出了可检测恶意攻击的隐私保护数据融合算法[5]。算法通过数据干扰有效保护了数据的隐私性，此外，树形结构的每一个子节点都会验证它的子孙节点是否恶意修改数据。Suat Ozdemir 等人提出了 SDFC[8]算法来实现隐私保护和完整性验证。为了防止节点注入虚假数据，算法中每个融合节点的所有监测节点对融合数据计算 MAC，从而对数据进行完整性验证。SDFC 算法的隐私保护采用的是逐跳加解密的方式，需要较大的计算开销。同时，SDFC 算法中部分节点是作为转发节点存在的，它们只是转发接收到的隐私数据而不会对数据进行完整性验证，这使得虚假数据不能被尽早地发现。

国内针对隐私保护也进行了大量的研究分析，陈娟等在文献[14]中针对 WSNs 中的安全问题及其攻防策略进行了深入讨论，并进一步指出隐私保护是 WSNs 安全领域未来的热点研究方向。文献[15,16]针对传感器网络中数据隐私保护问题的研究成果进行综述，介绍了传感器网络隐私保护网络模型、攻击模型和安全目标，阐述了代表性协议的关键实现技术。

在 SMART 的基础上，杨庚等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[17]。该算法依靠数据融合树型结构本身的特性，减少数据通信量；此外，算法随机分配时间片，以避免节点间的碰撞，并限制节点间串通的数据范围，降低数据丢失对精确度的影响。文献[18]提出了一种能同时保障数据融合技术 end-to-end 机密性与可认证性的安全数据融合协议。协议对密文信息进行了散列，并采用对称加密算法加密散列值，且不存在信息交互，因此较采用数字签名方案在计算量与通信量上都有明显改善。

3 系统模型

本文的无线传感器网络由一个基站节点和 N 个传感器节点组成，整个网络是以基站为根节点的树形结构，传感器节点采集到的数据沿着树形结构向上融合转发。每个传感器节点拥有唯一的 ID(从 1 到 N)，基站节点的 ID 为 0。我们假设基站节点知道整个传感网络的拓扑结构，并且拥有强大的计算、存储和通信能力；而非叶子节点知道它的子节点情况。每个传感器采集的数据都属于区间$[0,u_d]$。

我们假设基站节点是绝对安全可靠的，而任何的传感器节点都可以被捕获。传感器节点在被捕获以后，就会以不同的方式对传感器网络进行攻击。本文的研究只考虑以下两种攻击：(1)外部或者被捕获的传感器节点窃听传输数据包；(2)被捕获的传感器节点篡改融合结果，并设法使基站节点接受被篡改的数据。

我们定义融合函数 $y(t)=f(d_1(t),d_2(t),\cdots,d_N(t))$，$d_i(t)(i=1,\cdots,N)$表示 t 时刻节点 i 采集的数据。典型的数据融合函数包括：SUM、AVERAGE、COUNT、MAX、MIN 等。本文以求和函数 SUM 为研究对象，即 $y(t)=\sum_{i=1}^{N}d_i(t)$，其实 SUM 函数并不局限，这是由于上面提到的其他融合函数都可以归结为 SUM 函数。例如，由于 $\max(x_1,\cdots,x_N)=\lim_{k->\infty}(x_1^k+\cdots+x_N^k)^{1/k}$ 和 $\min(x_1,\cdots,x_N)=\lim_{k->-\infty}(x_1^k+\cdots+x_N^k)^{1/k}$，我们评估 max 和 min 时可以假定 k 是一个很大的值，这样 $\max(x_1,\cdots,x_N)\approx(x_1^k+\cdots+x_N^k)^{1/k}$、$\min(x_1,\cdots,x_N)\approx(x_1^k+\cdots+x_N^k)^{1/k}$。本文中引入的数据符号及其意义如表 1 所列。

表 1 PDI 的数据符号及意义

数据符号	意义
N	网络节点总数
A_u	当前节点
A_f	A_u 的父节点
A_{bi}	A_u 的第 i 个子节点
K^u_{group}	节点 u 与它的子节点共享的组密钥
n_u	A_u 的子节点数
n_f	A_f 的子节点数
T	预定义系统参数，监测节点数的上限值
D_u	当前节点 A_u 采集到的原始数据
$\tilde{D}_u$	对 D_u 进行数据扰动后得到的值
h(·)	安全单向散列函数
k_u	节点 u 和基站节点共享的密钥
$\tilde{D}_{aggi}$	节点 i 最终的融合结果
$subMAC_j(\tilde{D}_{aggi})$	节点 i 的第 j 个监测节点对融合结果计算的消息认证码
$FMAC(\tilde{D}_{aggi})$	节点 i 及其监测节点计算的所有消息认证码的串接值
$(M^k_{u,f},M'^k_{u,f})$	A_u 与 A_f 生成的第 k 个监测节点对

4 动态监测完整性验证算法

本文传感器网络采用树形结构，网络节点总数为 N，节点采集到的数据都属于区间$[0,u_d]$。基站节点向每一个传感器节点 u 预先载入共享密钥 k_u 以及单向散列函数 $h(\cdot)$。

基站有一个随机数池，每次发起查询时从中选取一个随机数 X_i；而且每次的随机数不重复，这保证了每次查询过程中提供数据隐藏功能的干扰数都不一样，有效提高了真实数

据的安全性。在树形结构中，基站节点沿着树向下发送查询信息$\langle query_statement, X_i\rangle$。

4.1 算法准备

4.1.1 组密钥建立

我们假设每个传感节点 A_u 和它的所有子节点创建一个组密钥 K^u_{group}，组密钥的创建采用现有的方案[9]；组密钥用于父节点和子节点间数据传输的完整性验证。

4.1.2 监测节点对形成

系统预定义监测节点数为 T，假设当前融合节点为 A_u，A_u 的子节点数为 n_u，A_u 的父节点 A_f 的子节点数为 n_f，则 A_u 与 A_f 之间实际监测节点对的数目为 $m=\min\{n_u, n_f-1, T\}$。

节点 A_u 将自己的子节点 ID 列表发送给 A_f，A_f 收到列表后按升序将 ID 进行排序并按序从 1 到 n_u 进行编号；同时，A_f 将自己的子节点(除 A_u 外)ID 列表进行升序排序并同样从 1 到 n_f-1 进行编号。然后 A_f 从编号 1 开始，找到 m 对编号相同的监测节点对。如图 1 所示，我们假设 T 为 3，则根据上述方法得：A_u 与 A_f 间形成了$\langle 2,4\rangle$，$\langle 3,5\rangle$两对监测节点对；类似地，A_{bi} 与 A_u 之间形成了$\langle 1,2\rangle$一对监测节点对。采用现有的随机密钥分配协议[10,11]来创建监测节点对间的共享密钥。

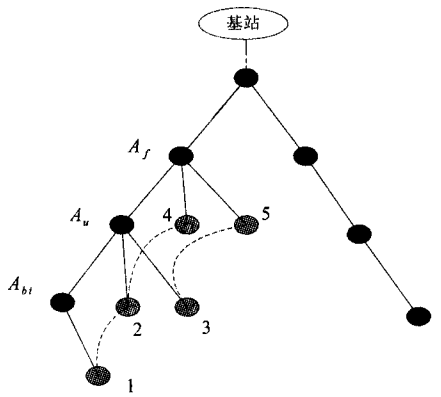

图 1 监测节点对

4.2 算法流程

4.2.1 隐私保护及数据融合阶段

假设当前节点为 A_u，A_u 的父节点为 A_f，A_u 的 n_u 个子节点为$\{A_{b1}, A_{b2}, \cdots, A_{bn_u}\}$。

为了保护数据的真实内容，A_u 对采集到的数据 D_u 进行扰动。基站有一个随机数池，每次发起查询时从中选取一个随机数 X_i，在树形结构中，基站节点沿着树向下发送查询信息$\langle query_statement, X_i\rangle$。每个节点收到 X_i 后计算 $P_u=h(k_u|X_i)$，k_u 是传感器节点 u 和基站节点之间的共享密钥，X_i 是只与当前查询相关的随机数，$h(\cdot)$是一个安全单向散列函数。得到 P_u 后，A_u 对采集到的数据 D_u 进行扰动，计算 $\tilde{D}_u=D_u+P_u \bmod q$($q=N\times u_d$，$u_d$ 是数据范围上限，N 是网络节点总数)。由于只有基站和节点知道 k_u，只有基站和节点才能计算出正确的 P_u 对数据进行还原，起到了隐私保护的作用。每次查询的随机数 X_i 不重复，这保证了每次查询过程中提供数据隐藏功能的干扰数都不一样，有效提高了真实数据的安全性。

在融合阶段，子节点 A_{bi}($1\leqslant i\leqslant n_u$)对融合数据 $\tilde{D}_{aggi}$ 计算 FMAC，在组内广播数据$\{\tilde{D}_{aggi}, FMAC(\tilde{D}_{aggi})\}$。当前节点 A_u 也同样在组内广播自己的数据 $\tilde{D}_u$，以便监测节点在完整性验证阶段进行验证。当融合节点 A_u 接收到来自 A_{bi}($1\leqslant i\leqslant n_u$)的数据$\{\tilde{D}_{aggi}, FMAC(\tilde{D}_{aggi})\}$时，$A_u$ 运用它与所有子节点(包括 A_{bi})共享的组密钥 K^u_{group} 对 $\tilde{D}_{aggi}$ 计算消息认证码，与接收到的 $FMAC(\tilde{D}_{aggi})$进行对比，验证数据是否被修改过。A_u 融合 $\tilde{D}_u$ 和所有 A_{bi} 发送过来的并通过完整性验证的数据 $\tilde{D}_{aggi}$($1\leqslant i\leqslant n_u$)，得到融合结果 $\tilde{D}_{aggu}$。$\tilde{D}_{aggu}=\tilde{D}_u+\sum_{i=1}^{n_u}\tilde{D}_{aggi}$ ($1\leqslant i\leqslant n_u$)。

4.2.2 数据还原阶段

在基站节点融合了所有发送过来的数据得到 $\tilde{D}_{agg0}$ 后，最终的真实融合结果为 $D_0=\tilde{D}_{agg0}-\sum_{j=1}^{S_u}P_j \bmod q$($S_u$ 是所有通过完整性验证的节点，$q=N\times u_d$，u_d 是数据范围上限，N 是网络节点总数)。由于基站和每个节点共享密钥 K_u，基站可以计算出 P_j 并得到去除干扰后的原始数据。

4.2.3 完整性验证阶段

如果 A_u 与 A_f 间存在监测节点对，由于所有的子节点和当前节点 A_u 在组内广播数据，监测节点同样能够融合所有数据得到融合结果 $\tilde{D}_{aggu}$。A_u 的监测节点 $M^k_{u,f}$ 使用与 A_f 的监测节点共享的密钥对融合结果 $\tilde{D}_{aggu}$ 计算消息认证码 $subMAC_k(\tilde{D}_{aggu})$($1\leqslant k\leqslant m$)并发送给 A_u。A_u 使用 A_f 与其子节点(包括 A_u)共享的组密钥 K^f_{group} 对融合结果 $\tilde{D}_{aggu}$ 计算消息认证码 $subMAC_u(\tilde{D}_{aggu})$，并与所有 $subMAC_k(\tilde{D}_{aggu})$($1\leqslant k\leqslant m$)串接成一个消息认证码 $FMAC(\tilde{D}_{aggu})$。最后，A_u 向上发送数据$\{\tilde{D}_{aggu}, FMAC(\tilde{D}_{aggu})\}$。

A_f 的监测节点 $M'^k_{u,f}$($1\leqslant k\leqslant m$)(与 A_u 的监测节点 $M^k_{u,f}$ 对应)接收到数据后，利用监测节点对之间共享的密钥同样对 $\tilde{D}_{aggu}$ 计算消息认证码，并与 $FMAC(\tilde{D}_{aggu})$进行对比进行完整性验证。一旦某一个监测节点完整性验证失败，A_f 就丢弃 $\tilde{D}_{aggu}$ 并通知基站节点。

PDI 的算法流程如图 2 所示。

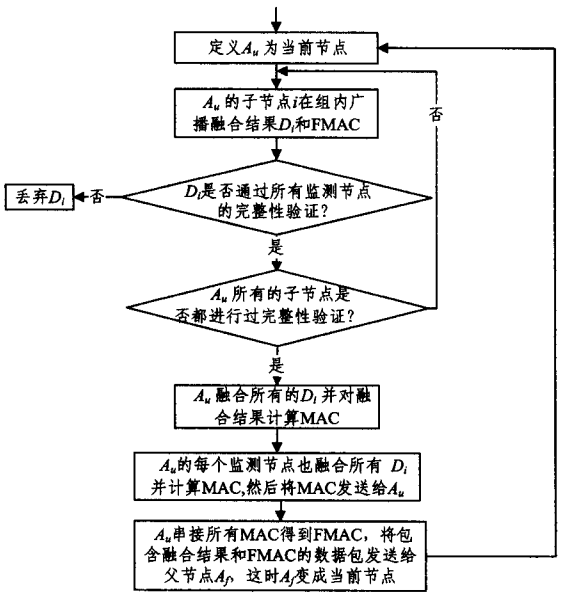

图 2 PDI 算法流程

5 性能分析

本节将主要从安全性、计算量、数据通信量、密钥分配这

4个方面分析本文所提PDI算法的性能。由于SDFC算法[8]是与PDI算法关系最为密切的算法，我们使用它作为以上4个方面的对比项。

本文采用TOSSIM软件对算法进行仿真，具体仿真环境为600个节点随机散落在$400\times400m^2$的区域中，基站节点位于网络的一角。节点传输范围为50m，高斯白噪声为4dB。尽管任何节点都可以感知事件并产生数据，仿真中我们假设主要是位于网络边缘的节点产生数据。

5.1 安全性分析

本文中，对算法的安全性分析主要是针对算法的虚假数据注入侦测能力进行分析。在算法SDFC中，被捕获节点可以在数据转发或者融合过程中篡改数据。当某个节点被捕获时，入侵者可以获得这个节点的所有敏感性息(如密钥等)。在被捕获节点不超过T个的情况下，算法SDFC可以侦测到注入的任何虚假数据。为了区分虚假数据和合法的融合数据，每个融合节点的所有监测节点都做数据融合并且为融合数据计算消息认证码。算法SDFC可以在数据融合过程中侦测到被捕获融合节点注入的虚假数据；同时，可以侦测到数据转发过程中注入的虚假数据。

本文提出的算法由于采用的是动态监测方案，每个融合节点的监测节点数不定，由当前节点A_u的子节点数n_u，A_u的父节点A_f的子节点数n_f，以及系统预定义监测节点数T决定。A_u与A_f之间实际监测节点对的数目为$m=\min\{n_u, n_f-1, T\}$，因此，本算法可以在被捕获节点不超过m的情况下检测到注入的虚假数据。

为了更直观说明相关性能，我们以实际应用为例。例如环境中利用100个传感器节点采集温度信息，并且采用PDI算法对采集到的温度数据进行隐私保护以及完整性验证。在安全参数T设为2的情况下，PDI算法产生大约24个拥有监测节点的融合节点，整个网络的监测节点总数为31个左右；而且随着传感器网络节点总数的增加，拥有监测节点的融合节点占网络节点总数的比率也会上升，安全性也会同步得到提高。

5.2 计算量分析

SDFC算法主要的计算开销发生在额外的消息认证码计算，数据融合节点和它的监测节点为明文数据和加密后的数据共计算$2\times(T+1)$个subMACs。同时，为了验证所有的subMACs，需要再计算$2\times(T+1)$个subMACs。除了计算消息认证码，数据融合节点需要解密和加密数据；T个监测节点需要解密广播数据并将所有数据融合。因此，SDFC算法总的计算开销包括：$4\times(T+1)$个MAC计算，$(T+1)$个数据融合处理，以及$(T+2)$个加解密过程。

本文所提算法由于采用干扰数的方式保护数据隐私，只需对干扰过的数据计算消息认证码，共需计算$(m+1)$个subMACs。为了验证完整性，同样需要再计算$(m+1)$个subMACs。m个监测节点需要进行数据扰动和数据融合操作，数据融合节点也要进行数据融合操作。因此，本文方案的总开销为：$2\times(m+1)$个MAC计算，$(m+1)$个数据扰动操作，以及$(m+1)$个数据融合处理。

对于计算量的仿真，我们主要估量针对侦测虚假数据需要的MAC计算。图3对比了网络中SDFC算法和PDI算法的MAC计算量。

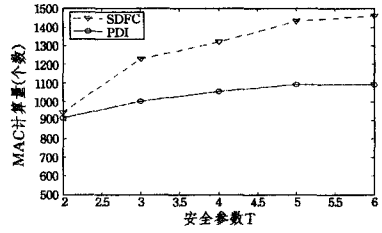

图3 SDFC和PDI的MAC计算量

由图3可以看出，当参数T增加时，SDFC算法和PDI算法的MAC计算量都相应增加。这是由于随着监测节点数的增加，两者都需要进行额外的MAC计算，而对于虚假数据注入的侦测能力也更强。同时，由于本文的PDI算法采用的是干扰数的方式进行隐私保护，不需要像SDFC算法一样分别对融合数据的密文和明文计算MAC，PDI算法的MAC计算量较小。

当T达到4的时候，算法PDI的MAC计算量增加平缓，这是由于PDI的监测节点数是由公式$m=\min\{n_u, n_f-1, T\}$(n_u为当前融合节点A_u的子节点数，n_f为A_f的子节点数)决定的。当T增加到一定程度时，网络中连续两个融合节点的子节点数都大于等于T的概率降低(即成T对监测节点对的概率降低)，这时MAC计算量增加减缓。

再次以100个传感器节点采集实际环境温度信息为例，将安全参数T设为2，则PDI算法的MAC计算量在106个左右。

5.3 通信量分析

由于算法SDFC使用两个4字节大小的FMACs，它的数据包大小表示为$(L_{tos}+4)$。L_{tos}表示TinySec[12]中一个认证加密数据包的长度。定义α为合法节点产生的数据包数，β为被捕获节点注入的虚假数据包数，H_d表示两个连续的融合节点间的平均跳数，H表示一个数据包在网络中传输的平均跳数。由于SDFC侦测两个连续融合节点间的虚假数据，虚假数据包最多可以传输H_d跳。SDFC的融合处理需要融合节点向T个监测节点广播每个数据，而T个监测节点要向融合节点发送T个subMACs。因此算法SDFC在整个传感器网络中的数据传输量为$D_{SDFC}=(L_{tos}+4)\times[(\alpha\times H)+(\beta\times H_d)]+T\times(L_{tos}+4)\times(\alpha+\beta)+T\times\frac{4}{T+1}\times(\alpha+\beta)$ bytes。

PDI算法只需对融合数据计算一个FMAC，它的数据包大小为L_{tos}。由于本文算法不存在单纯的转发节点，数据每向上传一次都会进行完整性验证，因此虚假数据在网络中只能传输1跳。本文算法融合处理同样需要融合节点向m个监测节点广播所有数据，而这m个监测节点要向融合节点发送m个subMACs。本文算法在整个传感器网络中的数据传输量为：

$$D_{PDI}=L_{tos}\times[(\alpha\times H)+\beta]+m\times L_{tos}\times(\alpha+\beta)+m\times\frac{4}{T+1}\times(\alpha+\beta)\text{ bytes}$$

仿真针对安全参数T对网络通信量开销的影响进行分析，纵坐标的数据通信量是网络中发送的数据包总个数，包括向上融合转发的数据包、为了验证需要融合节点发送给监测

节点的数据包、监测节点发送给融合节点的 MAC 数据包。从图 4 可以看出，当 T 增加时 SDFC 和 PDI 算法通信量都相应增加。尽管 T 的值不影响数据包的大小，但是它影响监测节点的个数，从而影响为了验证需要融合节点发送给监测节点的数据包以及监测节点发送给融合节点的 MAC 数据包。同时，图 4 显示 PDI 的通信量较 SDFC 小 5%左右，这是由于 PDI 父子节点间都进行完整性验证能更早地丢弃注入的虚假数据。此外，PDI 算法的监测节点只需传输一个 MAC 给融合节点。

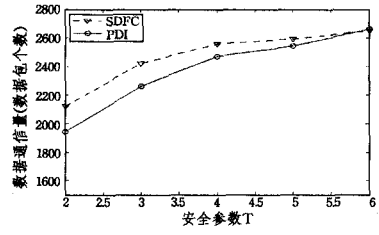

图 4　PDI 的数据通信量

仿真进一步针对虚假数据的注入量对 PDI 算法的数据通信量影响进行分析，如图 5 所示，横坐标代表虚假数据包数量和合法数据包数量的比值。随着比值的上升，SDFC 算法和 PDI 算法的数据通信量都有所上升，但是 PDI 的上升较 SDFC 平缓，这是由于 PDI 能比 SDFC 更早地丢弃注入的虚假数据，而 SDFC 可能会将虚假数据传输 H_d（连续两个融合节点的跳数）跳。

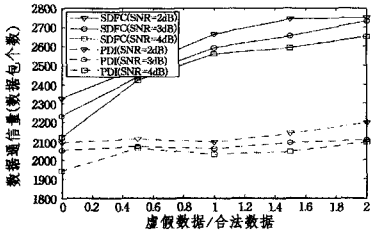

图 5　数据通信量

同时，图 5 显示了不同的 SNR 值对于两个算法的网络数据通信量的影响。由图可以看出，SDFC 受 SNR 的影响比 PDI 大，这是因为 SDFC 包的长度略大（有额外 4 个字节的 MAC），受丢包的影响更大。

同样地，在 100 个传感器节点采集温度信息的例子中，在安全参数 T 设为 2 的情况下，PDI 算法的数据通信量在 460 个数据包左右。

5.4　密钥分析

为了实现虚假数据侦测、安全数据融合以及隐私保护，SDFC 算法需要形成 $2T+1$ 对节点对，包括当前节点 A_u 的监测节点和 A_u 的转发节点间共享的 T 对密钥、当前节点 A_u 的监测节点和父节点 A_f 的邻居节点共享的 T 对密钥，以及 A_u 和 A_f 间共享的密钥。此外，融合节点 A_u 和它的邻居节点形成一个组密钥，组密钥的形成采用现有的方案[9]。

PDI 算法需要当前节点 A_u 的监测节点和父节点 A_f 的监测节点间共享的 m 对密钥，A_u 和它的邻居节点（即子节点）形成一个组密钥。算法实现传感数据隐私保护是基于干扰数的，干扰数的产生需要传感网络中每个节点都与基站节点共享一个密钥。

以利用 100 个传感器节点采集温度信息为例，PDI 算法的监测节点间共享的密钥对的个数就是监测节点对的个数，同样在 31 个左右。

结束语　作为 SDFC 算法的改进算法，PDI 算法不需要特定的结构，它根据现有的树形结构产生动态监测节点，而且每个节点都可以融合并且转发数据。PDI 算法使用数据扰动进行数据隐藏，所以不需要对传输的融合数据进行逐跳加解密。仿真结果显示，PDI 算法在计算量和通信量方面都低于 SDFC 算法。

PDI 算法实现了无线传感器网络中数据隐私保护和完整性验证的有机结合，能有效保障数据传输过程中的私密性和节点接收到的数据的有效性；而且，算法能够很好地利用网络带宽和传感器节点有限的能量。但是算法的实现需要在父节点及其子节点间共享组密钥，虽然组密钥的形成采用现有方案，但仍然给算法的实现带来了一定的复杂度。此外，算法的安全性完全取决于监测节点数 m，然而 m 的值可能为 0，即某些节点根本就没有监测节点，这使得算法在局部范围内安全性很低，数据容易被篡改。

由于 PDI 算法的监测节点数是动态生成的，因此每个融合节点的监测节点是不可预知的。如何避免出现监测节点数 m 为 0 的情况，进一步提高各个节点的安全性，这是下一步将要展开研究的方向。

参 考 文 献

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

[2] He Wenbo, Hoang Nguyen, Liu Xue, et al. Pda: Privacy-preserving data aggregation in wireless sensor networks[C]// Proceedings of 26th IEEE International Conference on Computer Communications (INFOCOMM). 2007: 2045-2053

[3] He Wenbo, Hoang Nguyen, Liu Xue, et al. iPDA: An Integrity-Protecting Private Data Aggregation Scheme for Wireless Sensor Networks[C]// Proceedings of IEEE Military Communications Conference. 2008: 1-7

[4] Feng T, Wang C, Zhang W, et al. Confidentiality protection for distributed sensor data aggregation[C]// Proceedings of 27th IEEE International Conference on Computer Communications (INFOCOMM). Phoenix, AZ, USA, 2008: 56-60

[5] Wang Chuang, Wang Gui-ling, Zhang Wen-sheng, et al. Reconciling Privacy Preservation and Intrusion Detection in Sensory Data Aggregation[C]// Proceedings of 27th IEEE International Conference on Computer Communications (INFOCOMM). 2011: 336-340

[6] Przydatek B, Perrig A, Song D. SIA: secure information aggregation in sensor networks[C]// Proceedings of the 1st international conference on Embedded networked sensor systems. New York, USA, 2003: 255-265

[7] Yang Y, Wang X, Zhu S, et al. SDAP: a secure hop-by-hop data aggregation protocol for sensor networks[J]. ACM Transactions on Information and System Security, 2008, 11(4)

（下转第 120 页）

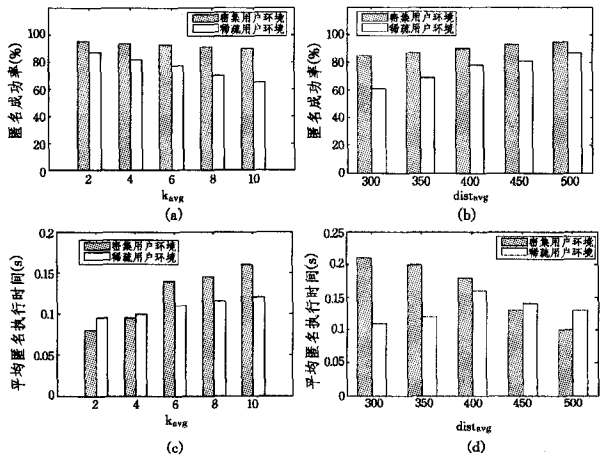

图 6 实验结果图

结束语 随着 LBS 在人们日常生活中的推广，位置隐私保护问题得到了广泛的讨论和研究。本文针对路网环境下的隐私保护特点，提出了基于路网 V 图的位置隐私保护方法，提出了新的位置隐私模型和匿名算法，用以对多个用户进行共同匿名，在保护用户隐私的同时提高匿名效率，保证服务质量。在今后的研究中，针对公路网络的结构特点，提出连续位置服务中的轨迹隐私安全保护方法，是下一步的研究目标。

参 考 文 献

[1] Freudiger J, Shokri R, Hubaux J-P. Evaluating the Privacy Risk of Location-Based Services[A]// 15th International Conference on Financial Cryptography and Data Security[C]. Gros Islet, St. Lucia, 2012: 31-46

[2] Cvrcek D, Kumpost M, Matyas V, et al. A study on the value of location privacy[A]// WPES '06: Proceedings of the 5th ACM workshop on Privacy in electronic society[C]. New York, USA: ACM, 2006: 109-118

[3] Gruteser M, Grunwald D. Anonymous usage of location-based services through spatial and temporal cloaking[A]// Proceedings of the 1st International Conference on Mobile Systems, Applications and Services[C]. San Francisco, USA: ACM, 2003: 163-168

[4] Sweeney L. k-anonymity: a model for protecting privacy[J]. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 2002, 10(5): 557-570

[5] Mokbel M F, Chow C Y. Casper*: query processing for location services without compromising privacy[J]. ACM Transactions on Data Systems, 2009, 34(4): 24-48

[6] Machanavajjhala A, Kifer D, Gehrke J, et al. l-diversity: Privacy beyond k-anonymity[J]. ACM Transactions on Knowledge Discovery from Data, 2007, 1(1): 3

[7] Gedik B, Liu L. Protecting location privacy with personalized k-anonymity: Architecture and algorithms[A]// IEEE Transactions on Mobile Computing[C]. California, USA, 2008: 1-18

[8] Shin K G, Ju X, Chen Z, et al. Privacy protection for users of location-based services[J]. IEEE Communications Society, 2012, 19(1): 30-39

[9] Kalnis P, Ghinita G, Mouratidis K, et al. Preventing location-based identity inference in anonymous spatial queries[J]. Knowledge and Data Engineering, 2007, 19(12): 1719-1733

[10] Wang Ting, Liu Ling. Privacy-aware mobile services over road networks[J]. VLDB Endowment, 2009, 2(1): 1042-1053

[11] 徐建，徐明，林欣. 路网限制环境中基于匿名蜂窝的位置隐私保护[J]. 浙江大学学报，2011，45(3)：429-434

[12] 薛姣，刘向宇，杨晓春. 一种面向公路网络的位置隐私保护方法[J]. 计算机学报，2011，34(5)：865-878

[13] 魏琼，卢炎生. 位置隐私保护技术研究进展[J]. 计算机科学，2008，35(9)：21-25

[14] Wu Xiao-jun, Luo Xue-fang. The Algorithm for Creating Weighted Voronoi Diagrams based on Cellular Automata[A]// The 6th World Congress on Intelligent Control and Automation[C]. Dalian China: Luo Xue-fang, 2006: 4630-4633

[15] Brinkhoff T. Network-based generator of moving objects[A]// Proceedings of the 12th International Conference on Scientific and Statistical Database Management[C]. Oldenburg, Germany, 2000: 253-255

(上接第 88 页)

[8] Ozdemir S, Çam H. Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks[J]. IEEE/ACM Transactions on Networking, 2010, 18(3): 736-749

[9] Blundo C, Santis A, Herzberg A, et al. Perfectly-secure key distribution for dynamic conferences[J]. Proceedings of Crypto, 1992, 740: 471-486

[10] Du W, Deng J, Han Y S, et al. A pairwise key pre-distribution scheme for wireless sensor networks[J]. ACM Transactions on Information and System Security, 2005, 8(2): 228-258

[11] Liu D, Ning P, Li R. Establishing pairwise keys in distributed sensor networks[J]. ACM Transactions on Information and System Security, 2005, 8(1): 41-77

[12] Karlof C, Sastry N, Wagner D. TinySec: A link layer security architecture for wireless sensor networks[C]// Proceedings of the 2nd international conference on Embedded networked sensor systems. New York, USA, 2004: 162-175

[13] Zhang W, Wang C, Feng T. Gp2s: Generic privacy-preservation solutions for approximate aggregation of sensor data[C]// Proceedings of IEEE PerCom. 2008: 179-184

[14] 陈娟，张宏莉. 无线传感器网络安全研究综述[J]. 哈尔滨工业大学学报，2011，43(7)：90-95

[15] 范永健，陈红，张晓莹. 无线传感器网络数据隐私保护技术[J]. 计算机学报，2012，35(6)：1131-1146

[16] 姚剑波，文光俊. 无线传感器网络中的隐私保护研究[J]. 计算机科学，2008，35(11)：19-22

[17] 杨庚，王安琪，陈正宇，等. 一种低耗能的数据融合隐私保护算法[J]. 计算机学报，2011，34(5)：792-800

[18] 张鹏，喻建平，刘宏伟. 传感器网络安全数据融合[J]. 计算机科学，2011，38(8)：106-108