# 无线传感器网络数据隐私保护技术

范永健[1),2),3)] 陈 红[1),2)] 张晓莹[1),2)]

[1)](中国人民大学数据工程与知识工程教育部重点实验室 北京 100872)

[2)](中国人民大学信息学院 北京 100872)

[3)](河北工程大学信息与电气工程学院 河北 邯郸 056038)

**摘 要** 研究和解决数据隐私保护问题对无线传感器网络的大规模应用具有重要意义,同时无线传感器网络的特征使得数据隐私保护技术面临严重挑战.目前无线传感器网络数据隐私保护技术已成为研究热点,主要针对数据聚集、数据查询和访问控制中数据隐私保护问题进行了研究.文中对无线传感器网络数据隐私保护现有研究成果进行了总结,从数据操作任务和隐私保护实现技术两个维度对现有研究成果进行了分类,介绍了网络模型、攻击模型和安全目标,阐述了代表性协议的关键实现技术,分析和比较了代表性协议的性能并总结了各协议的主要优缺点,最后指出了未来的研究方向.

**关键词** 物联网;无线传感器网络;隐私保护;数据聚集;数据查询;访问控制

**中图法分类号** TP309 **DOI号**: 10.3724/SP.J.1016.2012.01131

# Data Privacy Preservation in Wireless Sensor Networks

FAN Yong-Jian[1),2),3)] CHEN Hong[1),2)] ZHANG Xiao-Ying[1),2)]

[1)](*Key Laboratory of Data Engineering and Knowledge Engineering (Remin University of China) of Ministry of Education*, *Beijing* 100872)

[2)](*School of Information*, *Renmin University of China*, *Beijing* 100872)

[3)](*School of Information and Electrical Engineering*, *Hebei University of Engineering*, *Handan*, *Hebei* 056038)

**Abstract** Data privacy preservation is essential to widespread deployment of wireless sensor networks. However, the characteristics of wireless sensor networks make it face serious challenges. Data privacy preservation techniques in wireless sensor networks have attracted more and more attentions, and its research is mainly concentrated on data privacy preservation solutions for data aggregation, data query and access control in wireless sensor networks. This paper surveys the state of the art of data privacy preservation techniques in wireless sensor networks. Existing research results are classified according to two dimensions which are data manipulation tasks and privacy-preserving techniques. This paper introduces the network model, the attack model and security goal of each type of protocols, describes the key techniques of the important protocols, analyzes and compares the performance of these protocols, and summarizes the main advantages and disadvantages of these protocols. At last, suggestions for future research works are put forward.

**Keywords** Internet of Things; wireless sensor network; privacy preservation; data aggregation; data query; access control

---

收稿日期:2012-02-10;最终修改稿收到日期:2012-04-13.本课题得到国家自然科学基金项目(61070056,61075053)、核高基国家重大专项项目(2010ZX01042-001-002-002)资助.**范永健**,男,1978年生,博士研究生,讲师,中国计算机学会(CCF)会员,主要研究方向为无线传感器网络、隐私保护、数据库.E-mail:fanyj_ruc@ruc.edu.cn.**陈 红**(通信作者),女,1965年生,博士,教授,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为数据库、数据仓库、无线传感器网络.E-mail:chong@ruc.edu.cn.**张晓莹**,女,1987年生,博士研究生,主要研究方向为无线传感器网络、隐私保护.

## 1 引 言

无线传感器网络(以下简称为传感器网络)作为物联网的重要组成部分,在环境监测、医疗卫生、智能家居、国防军事等领域具有广阔的应用前景.随着传感器网络的应用发展,在实际应用部署过程中面临严重的隐私数据泄漏或被篡改的威胁.在医疗应用领域,使用便携式无线传感器对患者或体检者的心率、血压等重要体征数据进行收集分析时,这些敏感数据可能被泄漏;在智能家居领域,将无线传感器部署在居民家中,用于收集和统计区域内水、电和煤气的用量及分布情况,为市政单位和政府提供决策依据,同时攻击者可能获得这些数据并推测出居民是否在家等日常活动情况;在军事领域,无线传感器被部署在需要监测和侦查的重要区域,其收集和查询的数据往往携带重要情报信息,如果数据被泄漏或被篡改将带来严重威胁或决策失误.传感器网络中这些严重的隐私数据泄漏威胁,影响了传感器网络的应用发展.研究和解决传感器网络中的数据隐私保护问题,对传感器网络的大规模应用具有重要意义.

传感器网络具有资源受限、分布式、自组织、多跳和以数据为中心等特征,且往往部署在无人值守、不容易控制的复杂环境中.传感器网络中数据隐私保护主要面临以下挑战:(1)部署环境不容易控制.由于传感器网络往往部署在野外等无人值守的环境中,攻击者除可能通过链路层窃听获取敏感信息外,还可能俘获控制或伪造传感器节点进行窃取或篡改敏感信息;(2)资源受限.资源受限是传感器网络的重要特征,能量消耗量往往直接影响传感器网络的寿命.这使得传统网络中数据挖掘或数据发布中适用的隐私保护技术往往不能直接应用于传感器网络;(3)网络形态的多样性.传感器网络能够应用于监测、医疗、军事等多个领域,其网络形态和特征往往不同,可能存在不同的攻击模型和隐私保护需求,需要有针对性地设计隐私保护协议.

传感器网络中隐私保护技术主要可分为数据隐私保护技术和位置隐私保护技术.位置隐私保护技术主要针对攻击者通过对通信模式的监测分析,企图获知数据源或基站等重要目标位置的攻击方式,采用路由随机选取和虚假信息源等技术手段,隐藏真实的通信模式从而保护位置信息.数据隐私保护技术主要针对攻击者通过链路层窃听或俘获控制传感器节点,企图窃取或篡改隐私信息的攻击方式,主要采用扰动、匿名和加密等隐私保护技术,实现在不泄露隐私信息的情况下完成数据聚集、数据查询和访问控制等任务,与位置隐私保护技术存在很大差别.

传感器网络数据隐私保护是最近几年才引起广泛关注的新兴研究领域,目前已成为研究热点,已有很多重要研究成果.本文对该领域的主要研究成果进行了回顾与总结,从数据操作任务和隐私保护实现技术两个维度对现有研究成果进行了分类,阐述了代表性协议的关键实现技术,分析和比较了各协议的性能和主要优缺点,指出了未来的研究方向.

本文第 2 节介绍相关技术;第 3 节介绍研究分类和隐私保护协议性能评估指标;第 4 节介绍研究基于的模型;第 5 节到第 7 节分别对数据聚集、数据查询和访问控制中隐私保护技术进行阐述,并对其性能进行分析和比较;第 8 节指出未来研究方向.

## 2 相关技术

传感器网络数据隐私保护技术既需要隐私保护相关技术来防止隐私信息泄漏和被篡改,同时需要传感器网络数据管理相关技术来完成数据聚集、数据查询和访问控制等任务,并进行性能优化以减少能量消耗、时间延迟和数据丢失率.

目前隐私保护技术在数据库领域的应用主要集中在数据挖掘和匿名发布两个领域[1].综合现有的研究成果,隐私保护技术主要可分为 3 类:数据扰动技术[2-3]、数据加密技术[4-5]和数据匿名化技术[6-7].数据扰动技术使用加随机数、交换等技术对原始数据进行扰动,但需要保证处理后的数据能够满足相关应用需求;在数据加密技术中,安全多方计算(Secure Mutiparty Computation,SMC)[4-5]是目前研究热点之一;数据匿名化技术主要成果有 $k$-anonymity[6]、$l$-diversity、$t$-Closeness[7]等模型.数据挖掘和数据发布领域的隐私保护技术为传感器网络数据隐私保护技术提供了借鉴和支撑,但是往往由于能量消耗大或不适合分布式环境等因素不能直接应用于传感器网络.

传感器网络数据管理技术研究的重要目标为减少能量消耗,有很多研究分别从数据存储[8]、路由建立[9]和查询策略[10-12]等方面进行了性能优化.传感器网络中在不泄露隐私信息的情况下完成数据聚集、数据查询和访问控制等任务时,可以使用这些技

术对算法性能进行优化.

## 3　研究分类与性能评估

### 3.1　研究分类

目前，传感器网络数据隐私保护技术主要针对数据聚集、数据查询和访问控制中数据隐私保护问题进行研究. 数据聚集中隐私保护技术主要针对聚集节点可能被俘获时，研究在聚集节点不能获知感知数据的情况下实现数据聚集，并对聚集结果进行完整性验证，该方向研究成果相对较多；数据查询中隐私保护技术主要针对两层传感器网络中高资源节点可能被俘获时，研究在高资源节点不能获知感知数据和查询信息的情况下实现查询操作，并对查询结果进行完整性验证，目前该方向主要对范围查询、Top-$k$ 查询和基于类型查询中隐私保护问题进行研究，针对其它复杂查询中隐私保护问题的研究还很少；隐私访问控制主要研究在网络拥有者、传感器节点和其它用户不能获知用户身份和访问模式等信息的情况下实施访问控制，目前该方向的研究才刚起步.

根据传感器网络数据操作任务不同，隐私保护协议针对的网络模型、攻击模型和安全目标也往往不同. 同时，隐私保护协议采用的隐私保护技术和优化策略不同，其隐私保护能力、算法性能和结果精确度等性能指标也存在较大差异. 本文从数据操作任务和隐私保护实现技术两个维度对已有研究成果进行分类，首先按数据操作任务分为隐私保护数据聚集协议、隐私保护数据查询协议和隐私访问控制协议，其次根据隐私保护实现技术对每类协议进行进一步划分，表 1 给出分类情况.

**表 1　传感器网络数据隐私保护技术研究分类**

| 操作任务 | | 主要技术 | 代表协议 |
|---|---|---|---|
| 隐私保护数据聚集 | | 逐跳加密机制 | CPDA[13]、DADPP[14]、SMART[13]、iPDA[15]、ESPART[16] |
| | | 端到端加密机制 | AHE[17]、SP[18]、SIES[19]、CDA[20]、IPHCDA[21] |
| | | 非加密策略 | KIPDA[22]、GP$^2$S[23] |
| 隐私保护数据查询 | 范围查询 | 桶模式 | Sheng & Li 的方案[24]、Shi 等人的方案[25]、Zhang 等人的方案[26] |
| | | 前缀成员验证技术 | SafeQ[27] |
| | Top-$k$ 查询 | 扰动和安全比较技术 | SafeTQ[28] |
| | 基于类型查询 | 椭圆曲线多项式转换技术 | ElliPS[29] |
| 隐私访问控制 | | 盲签名 | DP$^2$AC[30] |
| | | 环签名 | Priccess[31] |

### 3.2　性能评估

目前传感器网络中数据隐私保护协议往往是针对特定数据操作设计的，仅支持特定的一种或几种数据操作，所以应考虑协议所支持的操作类型. 同时传感器网络资源的受限性和应用领域对聚集和查询结果精确度的需求，使得在关注隐私保护能力的同时应考虑算法性能和结果精确度. 本文从所支持操作类型、隐私保护能力、算法性能和结果精确度 4 个方面对协议性能进行评估与比较.

(1) 所支持操作类型. 所支持操作类型反映了协议的适用性和通用性.

(2) 隐私保护能力. 应考虑算法能够应对的攻击模型和隐私保护度. 在应对的攻击模型方面，考虑是否能够同时应对窃取信息攻击和篡改信息攻击以及应对共谋(collusion)攻击等能力；隐私保护度通常通过隐私信息披露风险来表示，披露风险越小，隐私保护度越高.

(3) 算法性能. 在算法性能方面主要考虑通信代价和计算代价，以评估能量消耗. 能量消耗往往直接影响传感器网络寿命. 同时，应根据服务需求考虑时间延迟、数据丢失率等指标.

(4) 结果精确度. 数据扰动等隐私保护技术往往影响结果的精确度，而数据聚集和查询结果的精确度影响基于此结果决策的准确性. 在保护数据隐私性的同时应考虑聚集和查询结果的精确度.

隐私保护技术往往需要以增加计算量和通信量为代价，同时隐私披漏风险往往需要与结果精确度进行折中，所以隐私保护协议设计应考虑隐私保护能力、算法性能、结果精确度等指标之间的平衡.

## 4　研究模型

### 4.1　攻击模型

对传感器网络隐私数据带来威胁的攻击方式可分为外部攻击和内部攻击两种模型. 由于传感器网络采用无线通信，数据在节点之间传输时攻击者可能通过链路层窃听获取敏感数据，这种攻击模式称为外部攻击，应对外部攻击主要采用数据加密和数据扰动等技术. 进行内部攻击时，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，能够获

取被俘获传感器节点的所有数据，并且能够获取密钥从而具有解密能力.因为内部攻击者具有一定的解密能力，单独应用逐跳加密方式(hop-by-hop encryption)将不再有效，传感器节点和基站共享密钥的端到端加密方式(end-to-end encryption)又给网内数据聚集和查询带来困难.这些挑战使得如何应对内部攻击成为数据隐私保护研究的主要内容.

根据攻击目的不同，内部攻击又可以分为窃取信息攻击和篡改信息攻击.窃取信息攻击模型又称为诚实但好奇模型(honest but curious threat model)[22]，被俘获的节点通过记录或推测敏感数据进行窃取信息，但其遵守协议且不篡改数据，仅破坏数据的隐私性.在篡改信息攻击模型中攻击者企图通过插入、修改或删除数据，使基站得到不正确或不完整的聚集或查询结果，破坏传感器网络中结果数据的完整性.在目前的研究成果中，隐私保护数据聚集和隐私访问控制研究中主要关注窃取信息攻击，隐私保护范围查询和隐私保护 Top-$k$ 查询研究同时考虑了窃取信息攻击和篡改信息攻击.

### 4.2 网络模型

传感器网络数据隐私保护研究主要基于节点相似的传感器网络和两层传感器网络两种网络模型.在节点相似的传感器网络中，因为网内节点在初始能量、存储能力、通信能力方面都是相似的，所以所有节点都有可能成为聚集节点等承担更多任务的节点，目前隐私保护数据聚集研究主要基于此网络模型.两层传感器网络由少量资源丰富的高资源节点(Master Node)和大量的资源受限的传感器节点(Sensor)组成，传感器节点负责采集感知数据，高资源节点负责收集数据和执行查询任务.文献[25-26，28]研究基于与图 1 相似的两层传感器网络模型.在此模型中，传感器节点之间构成下层网络，将数据传送至高资源节点，高资源节点之间能够进行长距离、高速率通信，形成多跳的上层网络，基站(Base Station)与一些高资源节点进行无线链接.两层传感器网络具有寿命长、易扩展等优点，使其具有良好的发展前景.目前，隐私保护数据查询研究主要基于两层传感器网络模型.

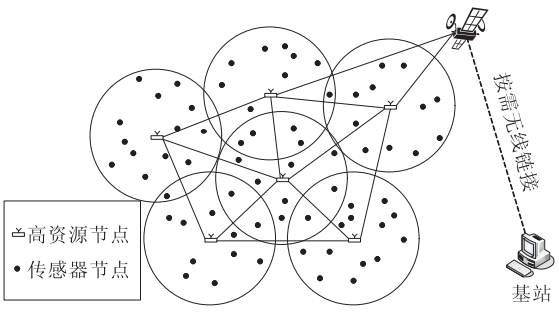


图 1 两层传感器网络模型示意图

## 5 隐私保护数据聚集

传感器网络数据聚集通过在聚集节点进行数据融合或压缩，能够有效地减少网络通信量，是传感器网络中重要的减少能量消耗技术.由于聚集节点需要收集聚集数据，容易成为攻击者攻击的对象.攻击者俘获控制聚集节点后，一方面能够获知经过聚集节点的明文数据，另一方面能够获得聚集节点密钥，能够对采用逐跳加密机制加密的数据进行解密，从而窃取或篡改所有经过聚集节点的敏感数据.隐私保护数据聚集的主要设计目标为在聚集节点不能获知所收集感知数据的情况下计算出聚集结果.采用传感器节点和基站共享密钥的端到端加密机制时，因为聚集节点不能进行解密，在聚集节点容易实现隐私保护，但是需要在加密数据上实现数据聚集.综合现有传感器网络隐私保护数据聚集研究成果，主要有 3 类实现策略：(1)使用逐跳加密机制.逐跳加密机制可以应对外部攻击，同时需要采用数据扰动(perturbation)、切分重组(shuffling)等数据失真(distorting)技术应对内部攻击；(2)使用端到端加密机制.需要使用同态加密模式在加密数据上实现数据聚集；(3)非加密策略.在不使用加密技术的情况下通过添加伪装数据、数据扰动等技术实现隐私保护数据聚集.

### 5.1 逐跳加密机制

在使用逐跳加密机制的协议中，聚集节点收到子节点上传的加密数据后，首先使用与其子节点共享的密钥进行解密，对所有解密的数据进行聚集，然后使用与其父节点共享的密钥对聚集结果进行加密并上传至其父节点.逐跳加密可以有效地应对外部攻击，但其将明文数据暴露给聚集节点，如果聚集节点被俘获控制将带来敏感信息泄漏风险.所以，在逐跳加密机制中需要使用数据扰动等其它隐私保护技术来应对内部攻击.数据扰动技术通过设计扰动模式使扰动后数据具有较高的隐私度，同时尽量减少或去除数据扰动对结果精确度的影响，保证恢复后的数据能够满足相关应用需求.由于在逐跳加密机制中每个中间节点都需要进行加密和解密操作，因此需要较高的计算代价和时间延迟.

5.1.1　扰动技术

CPDA(Cluster-based Privacy Data Aggregation)[13]中传感器节点通过在原始数据中添加随机种子和私有随机数进行扰动处理来隐藏真实数据值，簇头节点利用多项式的代数性质求解出精确的SUM聚集结果. 基本步骤为

(1)在概率参数控制下形成簇，簇由簇头节点和成员节点构成.

(2)簇内节点使用感知数据和本节点产生的随机数和随机种子数计算扰动数据，然后簇内节点使用两两共享密钥加密后交换扰动数据，各节点对所有收到的数据进行解密后执行求和操作，并将求和结果送至簇头节点，簇头节点根据聚集结果和种子数建立方程组，通过高斯消元法求解出精确的SUM聚集结果.

(3)簇头节点使用TAG(Tiny AGgregation service for ad hoc sensor networks)路由将计算出的聚集结果向基站传送.

设CPDA形成的簇内有 $n$ 个成员节点，$s_0$ 表示簇头节点，$s_1,\cdots,s_n$ 分别表示成员节点，$v_i$ 表示节点 $s_i$ 采集的数据. 簇内每个节点 $s_i(0\leq i\leq n)$ 产生非零种子 $x_i$，将 $x_i$ 在簇内进行广播使所有节点共享种子，同时，$s_i(0\leq i\leq n)$ 产生 $n$ 个私有随机数 $r_1^i,\cdots,r_n^i$ 并计算扰动数据 $V_j^i=v_i+r_1^i x_j+r_2^i(x_j)^2+\cdots+r_n^i(x_j)^n$，其中 $0\leq j\leq n$，然后，$s_i$ 使用与 $s_j$ 共享密钥 $k_{ij}$ 对 $V_j^i$ 加密后发送至 $s_j$. 在簇内节点两两交换扰动数据后，节点 $s_j(0\leq j\leq n)$ 对所有收到的扰动数据解密后进行求和得到组合数据 $F_j=\sum_{i=0}^{n}V_j^i=v_{\text{sum}}+r_1x_j+r_2(x_j)^2+\cdots+r_n(x_j)^n$，其中 $v_{\text{sum}}=\sum_{i=0}^{n}v_i$，$r_k=\sum_{i=0}^{n}r_k^i$ $(1\leq k\leq n)$，$s_j$ 将组合数据 $F_j$ 广播至簇头节点 $s_0$. 此时，簇头节点 $s_0$ 得到 $n+1$ 个方程 $F_j(0\leq j\leq n)$ 组成的以 $\boldsymbol{G}$ 为系数矩阵的线性方程组，其中 $\boldsymbol{G}=\{1,\boldsymbol{X},\cdots,\boldsymbol{X}^n\}$ $(\boldsymbol{X}=\{x_0,x_1,\cdots,x_n\}^{\mathrm{T}})$，因为 $\boldsymbol{X}=\{x_0,x_1,\cdots,x_n\}^{\mathrm{T}}$ 中数据项互不相同，所以 $\boldsymbol{G}$ 为满秩矩阵，$s_0$ 由 $\boldsymbol{U}=\boldsymbol{G}^{-1}\boldsymbol{F}$ 可以求解出向量 $\boldsymbol{U}=\{v_{\text{sum}},r_1,\cdots,r_n\}^{\mathrm{T}}$，其中 $\boldsymbol{F}=\{F_0,F_1,\cdots,F_n\}^{\mathrm{T}}$. 这样，簇头节点 $s_0$ 在不能获知其它节点真实数据值的情况下，求解出了精确SUM聚集结果 $v_{\text{sum}}$. 图2演示了 $n=2$ 时CPDA描述的信息交换过程.

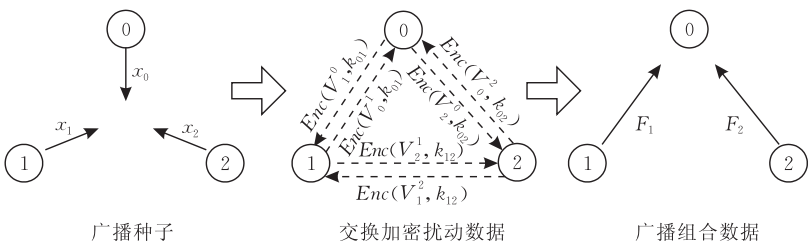


图2　CPDA信息交换过程演示

DADPP(Data Aggregation Different Privacy-levels Protection)[14]能够提供不同隐私保护水平的聚集处理，其处理过程类似于CPDA. DADPP按不同隐私保护水平将簇内传感器节点分成若干组(group)，组内节点具有相同的隐私保护水平. 首先对组内数据进行预处理并将结果上传给簇头节点，簇头节点根据预处理结果计算出聚集结果并上传.

5.1.2　切分重组技术

SMART(Slice-Mixed AggRegaTion)[13]使用切分重组技术完成隐私保护数据聚集. SMART基本思路为：传感器节点将原始数据随机地切分为数个数据切片(pieces)，采用逐跳加密机制与随机选择的邻居节点交换数据切片，对所有收到的数据切片执行求和操作，并将求和结果上传至基站，基站对所有收到的数据进行求和，得到精确的SUM聚集结果. SMART分为3个步骤.

Slicing. 网络中每个节点 $s_i(i=1,2,\cdots,N)$ 在 $h$ 跳内随机选择 $J-1$ 个邻居节点构成节点集 $S_i$，将感知数据 $d_i$ 随机切分为 $J$ 个数据切片，$s_i$ 为本节点留下其中1个数据切片，将其余 $J-1$ 个数据切片采用逐跳加密机制加密后分别随机发送至 $S_i$ 中节点，用 $d_{ij}$ 表示由节点 $s_i$ 传送至 $s_j$ 的数据切片.

Mixing. 每个节点 $s_j$ 对收到的数据解密后求和得到 $r_j=\sum_{i=1}^{N}d_{ij}$，其中当 $j\notin S_i$ 时 $d_{ij}=0$.

Aggregation. 所有节点 $s_j(j=1,2,\cdots,N)$ 使用树形路由(tree-based routing)将计算的 $r_j$ 送至基站，基站对所有 $r_j$ 求和得到求和聚集结果 $\sum_{j=1}^{N}r_j$. 图3演示了 $N=5$，$J=2$，$h=1$ 时SMART数据切片与交换情况.

iPDA(integrity-Protecting private Data Aggre-

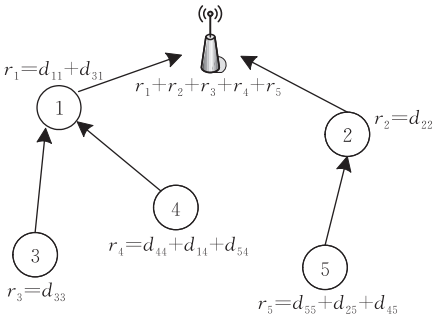


图 3 SMART 数据切片与交换过程演示

gation)[15]使用 SMART 中数据切分重组思路实现数据聚集中隐私保护,并通过冗余数据对聚集结果进行完整性验证.为了验证完整性,iPDA 建立两棵节点无交集的聚集树,每个节点需要将感知数据传至两棵聚集树,两棵聚集树独立进行数据聚集,基站通过比较两棵聚集树的聚集结果,可以验证聚集结果的完整性,如果相等则判断为数据完整,否则判断为数据被篡改或部分节点失效.冗余数据需要额外通信量.

ESPART (Energy-Saving Private-preserving AggRegaTion)[16]对 SMART 进行了改进,在节省能量消耗的情况下,实现了隐私保护数据聚集.ESPART 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,并限制了节点间共谋(collusion)的数据范围,降低了数据丢失对精确度的影响,在精确度要求相近的情况下,需要较少的数据融合时间.ESPART 依靠数据融合树形结构本身的特性,减少了通信量和能量消耗.

### 5.2 端到端加密机制

端到端加密机制使用传感器节点和基站共享的密钥加密,聚集节点不能解密,能够较好地应对内部攻击和外部攻击.与逐跳加密机制相比,端到端加密机制中中间节点节省了加解密计算代价,减少了时间延迟.端到端加密机制需要在加密数据上实现数据聚集.使用同态加密模式(Homomorphic Encryption scheme)可以实现在密文上进行求和或乘积操作,能够有效地支持在加密数据上实现数据聚集.用 $m_1$ 和 $m_2$ 表示两个明文数据,用 $\odot$ 表示求和或乘积操作.如果加密函数 $Enc()$ 为同态加密函数,则能够通过 $Enc(m_1)$ 和 $Enc(m_2)$ 直接运算得到 $Enc(m_1\odot m_2)$,通过解密函数 $Dec()$ 对 $Enc(m_1\odot m_2)$ 解密能够得到明文求和或乘积结果即 $Dec(Enc(m_1\odot m_2))=m_1\odot m_2$.

AHE(Additively Homomorphic Encryption)[17]使用一种简单的求和同态加密函数,实现采用端到端加密机制的 SUM 聚集.AHE 中加密和解密函数分别为

$$Enc(m,k,M)=m+k \bmod M,$$
$$Dec(c,k,M)=c-k \bmod M.$$

传感器节点 $s_i$ 与基站共享密钥 $k_i$,$k$ 由 $k_i$ 计算得到.$M$ 为系统参数,用 $m_1$,$m_2$ 分别表示传感器节点 $s_1$,$s_2$ 采集的明文数据,AHE 实现聚集如下:

传感器节点 $s_1$:$c_1=Enc(m_1,k_1,M)$;

传感器节点 $s_2$:$c_2=Enc(m_2,k_2,M)$;

聚集节点:$c_{12}=(c_1+c_2) \bmod M$;

基站:$Dec(c_{12},k_1+k_2,M)=m_1+m_2$.

AHE 中基站解密时需要上传与上传数据对应的传感器节点 ID,以确定解密函数中参数 $k$,所以传感器节点上传加密数据时需要上传节点 ID,这样需要额外的通信开销.AHE 不支持聚集结果完整性验证.

和 AHE 思路相似,文献[18]提出 SP(Secret Perturbation)模式系列.其中 BSP 模式通过添加辅助数据项,在实现 AHE 功能的同时能够应对重放攻击(replay attacks),提高了隐私保护度;FSP 模式要求不需要上传感知数据的节点发送秘密扰动数,这样所有节点都上传了数据,基站将收到的数据减去所有节点的秘密扰动数之和,即得到正确的聚集结果,FSP 中传感器节点不再需要上传节点 ID;O-ASP模式和 D-ASP 模式对 FSP 模式进行了优化以减少通信量.文献[19]提出 SIES(Secure In-network processing of Exact SUM queries)模式,使用求和同态加密函数[32]和共享秘密数据[33]等技术在密文上完成数据聚集,采用 SECOA[34]思路实现聚集结果完整性验证.

CDA (Concealed Data Aggregation)[20]采用 Domingo-Ferrer 提出的同态加密模式(DF scheme)实现传感器网络中隐私保护数据聚集.DF scheme 将数据 $a$ 切分成 $m(m\geq 2)$ 项 $\{a_1,a_2,\cdots,a_m\}$,有 $a=\sum_{j=1}^{m}a_j$,设密钥 $k=(r,g')$,使用以下加密函数对 $a$ 进行加密

$Enc_k(a)=\{a_1 r \bmod g, a_2 r^2 \bmod g,\cdots,a_m r^m \bmod g\}$,其中 $g$ 与 $g'$ 相关.

解密时对于第 $j$ 项使用 $r^{-j} \bmod g$ 计算 $a_j \bmod g$,则解密函数可以实现

$$Dec_k(Enc_k(a))=\sum_{j=1}^{m}a_j \bmod g'.$$

CDA 中所有传感器节点与基站共享密钥 $k=(r,g')$.传感器节点 $s_i$ 将感知数据 $d_i$ 加密得 $Enc_k(d_i)$,并上传至聚集节点 $A$;设聚集节点 $A$ 收到

$n$ 个加密数据，$A$ 对所有收到的加密数据进行聚集得到 $y'=f(E_k(d_1),\cdots,E_k(d_n))$，将 $y'$ 上传至基站；基站解密得聚集结果 $y=Dec_k(y')$.

IPHCDA（Integrity Protecting Hierarchical Concealed Data Aggregation）[21] 采用基于椭圆曲线的同态加密模式[35]和消息认证码（message authentication codes）来应对窃取信息攻击和篡改信息攻击，实现分层数据聚集中的隐私保护和完整性验证. 在网络部署阶段，IPHCDA 将网络划分为若干个分区. 在数据聚集时，不同的分区中传感器节点使用不同的公钥对采集数据进行加密，并计算聚集数据的消息认证码，将加密数据和消息认证码上传至聚集节点；聚集节点对加密数据进行聚集后上传聚集结果；基站对数据解密时，能够根据使用密钥的不同对加密数据按分区进行分类，并验证聚集数据的消息认证码.

### 5.3　非加密策略

KIPDA（K-Indistinguishable Privacy-preserving Data Aggregation）[22] 在不加密的情况下通过添加伪装数据（camouflage data）实现隐私保护 MAX/MIN 非线性聚集，并可扩展实现隐私保护 SUM 聚集. KIPDA 基本思路为：传感器节点将采集的真实数据和伪装数据构成消息集，消息集中数据的位置是专门安排的，传感器节点向聚集节点上传消息集，因为真实数据没有被加密，聚集节点能够实现数据聚集，同时对于攻击者真实数据与伪装数据具有不可区分性，能够在不加密的情况下实现隐私保护非线性聚集.

用 $U^i=\{v_1^i,v_2^i,\cdots,v_n^i\}$ 表示节点 $i$ 的消息集，用 $I=\{1,2,\cdots,n\}$ 表示 $U^i$ 位置索引集，有 $|I|=|U^i|$. KIPDA 分为预配置、报告、聚集、基站处理 4 个阶段.

（1）在预配置阶段，基站进行如下工作：选择 $I$ 的子集作为全局秘密信息集 $GSS$，有 $GSS\subset I$，$GSS$ 为基站私有秘密信息集；为每个节点 $s_i$ 选择 $GSS$ 的子集 $NSS_T^i$ 用于存放真实的采集数据，有 $NSS_T^i\subset GSS$，MAX/MIN 聚集时 $|NSS_T^i|=1$；为每个节点 $s_i$ 确定节点秘密信息集 $NSS^i$，$NSS^i$ 由 $GSS$ 和 $\overline{GSS}$ 的子集组成，有 $GSS\subset NSS^i$ 且 $\overline{GSS}\cap NSS^i\neq\varnothing$，$NSS^i$ 由基站和节点 $i$ 共享秘密信息集；将 $NSS_T^i$ 和 $NSS^i$ 下发至相应节点 $i$. 用图 4 示例说明，假设基站确定 $|I|=7$、$|GSS|=3$、$|NSS^i|=5$，令 $GSS=\{2,4,6\}$，则 $\overline{GSS}=\{1,3,5,7\}$，按规则令 $NSS_T^1=\{2\}$、$NSS_T^2=\{6\}$、$NSS_T^3=\{4\}$、$NSS^1=\{2,4,6,1,3\}$、$NSS^2=\{2,4,6,3,7\}$、$NSS^3=\{2,4,6,1,7\}$.

（2）在报告阶段，每个节点 $i$ 生成消息集 $U^i$ 并传送至聚集节点. 用 $[d_{\min},d_{\max}]$ 表示感知数据的取值范围. 在 $NSS_T^i$ 对应的 $U^i$ 的位置，放入真实的感知数据 $d_i$；在 $NSS^i-NSS_T^i$ 对应的 $U^i$ 的位置填充满足以下条件的伪装数据：当 MAX 聚集时，伪装数据范围为 $[d_{\min},d_i]$，当 MIN 聚集时，伪装数据范围为 $[d_i,d_{\max}]$；在 $\overline{NSS^i}$ 对应的 $U^i$ 的位置填充范围为 $[d_{\min},d_{\max}]$ 的伪装数据. 图 4 示例中执行 MAX 聚集，感知数据的范围为 $[0,30]$，$N1$、$N2$ 和 $N3$ 节点采集的数据值分别为 16、28 和 12. 按规则随机生成伪装数据并填充后得到 $U^1=\{2,16,15,8,23,6,16\}$、$U^2=\{29,22,16,5,9,28,26\}$、$U^3=\{8,1,18,12,21,10,3\}$.

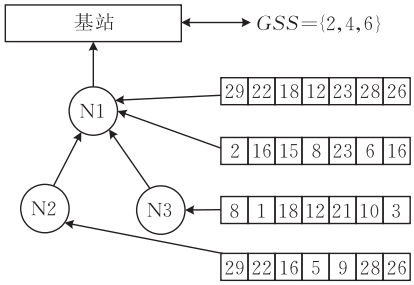


图 4　KIPDA 数据聚集演示

（3）在聚集阶段，聚集节点 $i$ 计算新的 $U^i$ 并上传. 用节点 $j$ 表示聚集节点 $i$ 的孩子节点. 聚集后新的 $U^i=\{v_1^i,v_2^i,\cdots,v_n^i\}$ 中数据项 $v_l^i(l=1,2,\cdots,n)$，为原 $U^i$ 与所有孩子节点上传的 $U^j$ 中相应 $l$ 位置数据的最值（MAX 聚集为最大值，MIN 聚集为最小值），即 $v_l^i=\max(\min)(v_l^j,v_l^i)$. 图 4 示例中计算的新的 $U^1$ 为 $\{29,22,18,12,23,28,26\}$.

（4）在基站处理阶段，基站得到最终聚集消息集 $U^f$，取 $U^f$ 中 $GSS$ 对应位置数据的最值（MAX 聚集为最大值，MIN 聚集为最小值），得到最终聚集结果，最终结果为 $\max(\min)_{k\in GSS}(v_k^f)$. 图 4 示例中最终结果为 $\max(22,12,28)$，即为 28.

在实现聚集 MAX（或 MIN）时，限制集 $NSS^i$ 对应 $U^i$ 中数据项均小于等于（或大于等于）真实数据 $d_i$，而非限制集 $\overline{NSS^i}\subset\overline{GSS}$ 不影响聚集结果，所以 KIPDA 能得到精确聚集结果. 同时，对于任何节点，KIPDA 可以实现 $k$ 匿名，其中 $k=\overline{NSS^i}+1$.

GP$^2$S（Generic Privacy-Preservation Solutions for approximate aggregation）[23] 采用扰动直方图聚集模式（perturbed histogram-based aggregation schemes），实现隐私保护数据聚集. 基本思路为：传感器节点将采集数据映射到直方图区间以实现数据的泛化处理，并将处理后数据加入传感器节点和基站共享的秘密数进行扰动处理，然后上传至聚集节

点进行求和聚集，聚集节点将聚集结果向基站方向传送，基站将收到的聚集结果减去扰动数据，经计算得到表示所有感知数据分布情况的直方图，使用此直方图能够计算出近似 MAX/MIN、SUM、AVERAGE、Median、Histogram 等聚集结果. 扰动和泛化技术使聚集节点不能获得真实感知数据和直方图.

传感器网络中有 $N$ 个传感器节点，每个传感器节点 $s_u$ 预装与基站共享的秘密数 $S_u$ 和单路哈希函数 $h(\cdot)$. 基站将采集数据值域 $[v_{\min}, v_{\max}]$ 均匀划分为 $n$ 个区间 $(D_1, D_2, \cdots, D_n)$，每个区间宽度为 $\sigma=(v_{\max}-v_{\min})/n$，则 $D_i$ 为 $[(i-1)\times\sigma, i\times\sigma]$. GP$^2$S 基本步骤如下：基站下发聚集查询 $\langle Query, \sigma, X\rangle$，其中 $X$ 标识本次聚集查询用于应对重放攻击. 接到 $\langle Query, \sigma, X\rangle$ 后，传感器节点 $s_u$ 根据区间 $(D_1, D_2, \cdots, D_n)$ 构造对应向量 $(d_{u,1}, d_{u,2}, \cdots, d_{u,n})$. 根据传感器节点是否为叶子节点进行不同处理. (1) 叶子节点 $s_u$ 进行如下处理：如果其采集数据 $v\in D_i$ 则 $d_{u,i}=[1+h(S_u|X|i)] \bmod N$；对于所有 $j(j\neq i, v\notin D_j)$，则 $d_{u,j}=h(S_u|X|j) \bmod N$；(2) 非叶子节点 $s_u$ 接收到全部 $m_u$ 个子节点上传的数据后进行如下处理：如果其采集数据 $v\in D_i$，则

$$d_{u,i}=\left[1+h(S_u|X|i)+\sum_{j=1}^{m_u} d_{j,i}\right] \bmod N;$$

对于所有 $j(j\neq i, v\notin D_j)$，则

$$d_{u,j}=\left[h(S_u|X|j)+\sum_{j=1}^{m_u} d_{j,i}\right] \bmod N.$$

将计算后的向量 $(d_{u,1}, d_{u,2}, \cdots, d_{u,n})$ 上传. 基站接收到其全部 $m_0$ 个子节点上传的数据后计算 $\left[\sum_{j=1}^{m_0} d_{j,i}-\sum_{u=1}^{N} h(S_u|X|i)\right] \bmod N$，得到直方图，然后根据直方图计算近似聚集结果.

### 5.4 性能分析与比较

采用逐跳加密机制时，中间节点需要加密和解密操作，需要较高的计算代价和时间延迟，逐跳加密机制能够应对外部攻击，应对内部攻击则需要使用扰动等其它技术；采用端到端加密能够较好应对内部和外部攻击，计算代价和时间延时较小，需要同态加密模式在加密数据上实现数据聚集，但该机制不能很好地完成 MAX/MIN 等非线性聚集操作；非加密策略不需要密钥分配和加解密操作，节省了计算和通信代价，但是需要添加伪装数据或数据扰动代价等技术以实现隐私保护，其性能取决于采用的隐私保护技术.

在支持聚集类型方面，CPDA、DADPP、SMART、iPDA、ESPART 和端到端加密模式仅支持 SUM 聚集，KIPDA 支持 MAX/MIN 和 SUM 等聚集类型，GP$^2$S 支持 MAX/MIN、SUM、AVERAGE、Median、Histogram 等聚集类型；在应对篡改信息攻击方面，iPDA、SIES 和 IPHCDA 提供应对篡改信息攻击的完整性验证机制，本节所述其它协议仅考虑外部攻击和窃取信息攻击；在应对传感器节点被俘获发起共谋攻击方面，GP$^2$S、AHE、SIES 和 SP 等协议采用每个传感器节点与基站共享秘密数或密钥机制，所以应对共谋攻击能力较强，CPDA、DADPP、SMART、iPDA、ESPART 和 KIPDA 具有一定应对共谋攻击能力，CDA 因为所有传感器节点与基站共享同一密钥，所以应对共谋攻击能力较弱；在隐私保护度方面，KIPDA 仅实现了 $k$ 个数据不可区分，隐私保护度相对于其它协议较弱；在算法性能方面，CPDA、DADPP 需要较高计算和通信代价，SMART、iPDA 通信代价较高而计算代价相对较小，KIPDA 和 GP$^2$S 节省了中间节点的加密和解密计算代价，同时由于需要由向量代替数据项进行上传，增加了相应通信量和存储代价，端到端加密机制需要的计算代价和时间延迟较小；在聚集结果精确度方面，GP$^2$S 提供近似聚集结果，本节所述其它协议能够提供精确聚集结果. 表 2 对隐私保护数据聚集代表性协议在性能方面的主要优缺点进行了总结.

**表 2 隐私保护数据聚集协议对比**

| 代表协议 | 采用技术 | | 主要优点 | 主要缺点 |
|---|---|---|---|---|
| CPDA[13]、DADPP[14] | 逐跳加密机制 | 扰动技术 | 精确聚集；DADPP 实现多隐私水平聚集 | 仅支持 SUM；通信和计算代价较高 |
| SMART[13]、iPDA[15]、ESPART[16] | | 切分重组 | 精确聚集；iPDA 提供完整性验证 | 仅支持 SUM；通信代价较高 |
| AHE[17]、SP[18]、SIES[19] | 端到端加密机制 | 同态加密函数 AHS | 应对共谋攻击能力强；SIES 提供完整性验证 | 仅支持 SUM |
| CDA[20] | | 同态加密函数 DF scheme | 通信和计算代价较小 | 仅支持 SUM；应对共谋攻击能力较弱 |
| IPHCDA[21] | | 基于椭圆曲线的同态加密函数 | 提供完整性验证；能够对加密数据分类 | 仅支持 SUM |
| KIPDA[22] | 非加密策略 | 伪装数据 | 支持非线性聚集 | 隐私保护度较弱 |
| GP$^2$S[23] | | 扰动直方图 | 支持多种类型聚集；应对共谋攻击能力强 | 近似聚集 |

## 6　隐私保护数据查询

目前传感器网络隐私保护数据查询主要基于两层传感器网络进行研究. 两层传感器网络中查询过程如图 5 所示，基站收到用户查询后，将查询转化为多个查询，通过上层网络将查询传输至相应高资源节点，由高资源节点完成相应查询，并将查询结果传输到基站，由基站汇总后返回给用户. 两层传感器网络有两类常用数据存储模式：(1) 高资源节点作为存储节点，附近节点将采集的数据存入存储节点，当高资源节点收到基站下发的查询后，从存储的数据中匹配符合查询要求的感知数据并返回查询结果；(2) 感知数据存储于传感器节点中，高资源节点收到查询后从传感器节点收集数据完成查询.

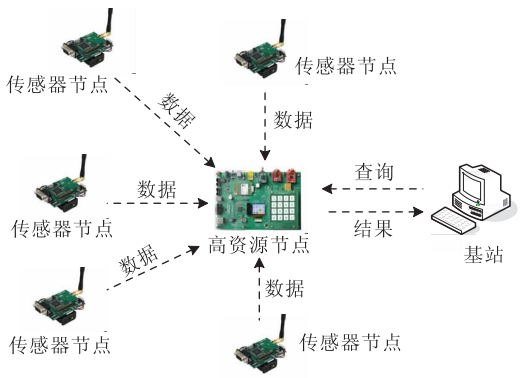


图 5　两层传感器网络中查询过程演示

在两层传感器网络中，高资源节点需要完成数据收集和查询任务，攻击者俘获高资源节点将对传感器网络安全造成严重的威胁. 攻击者俘获高资源节点后，可能采用窃取信息和篡改信息两种攻击模型，破坏数据的隐私性和完整性.

在查询类型方面，目前主要针对范围查询、Top-$k$ 查询和基于类型查询中隐私保护问题进行研究.

### 6.1　隐私保护范围查询

文献[24-27]对隐私保护范围查询进行研究，基于的网络模型为高资源节点为存储节点的两层传感器网络；攻击模型为攻击者通过俘获存储节点进行窃取信息攻击和篡改信息攻击；安全目标为在存储节点不能获得感知数据和查询数据的情况下，完成范围查询并且能够检测不正确和不完整查询结果. 上述文献所采用的主要技术为桶模式和前缀成员验证技术.

#### 6.1.1　基于桶模式协议

Sheng & Li 的方案[24]首次考虑传感器网络范围查询中隐私保护问题[24]，采用桶模式(bucketing scheme)[36-37]和加密技术完成隐私保护范围查询，通过添加验证编码进行完整性验证. 基本思路为：传感器节点和基站保持同样的桶划分，传感器节点将采集的数据映射到相应的桶中，使用与基站共享密钥按桶分别进行加密，附加桶标签(Tag)后上传至存储节点进行存储，基站将范围查询对应的标签发给存储节点，存储节点进行标签匹配后，将相应加密数据返回，基站解密后得查询结果. 具体实现过程如下：系统将感知数据值域划分为多个无交叉且连续的桶(buckets)，这些桶都有唯一的标签. 基站和传感器节点都知道桶划分方式. 在时间段 $t$ 传感器节点 $s_i$ 将采集的感知数据存入相应的桶. 在时间段 $t$ 结束时，传感器节点 $s_i$ 使用与基站共享密钥 $k_{i,t}$ 对每个桶中数据整体加密. 对于没有数据的桶 $T_j$，生成验证编码 $num(i,j,t)$ 作为查询结果完整性验证信息，其中 $num(i,j,t)=H_j(i \| j \| t \| k_{i,t})$. 示例如下，假设桶划分数为 3，在时间段 $t$ 传感器节点 $s_i$ 在桶 $T_1$，$T_3$ 分别有 1，2 项数据，传感器节点 $s_i$ 传送下面信息到存储节点(Storage Node)：

$$
\begin{aligned}
&s_i \rightarrow Storage\ Node: i,t,\\
&\langle T_1,(data_1)_{k_{i,t}}\rangle,\\
&\langle T_2,num(i,2,t)\rangle,\\
&\langle T_3,(data_2,data_3)_{k_{i,t}}\rangle
\end{aligned}
$$

当用户向基站提交范围查询$\{t,[a,b]\}$后，基站将此范围查询转换为能够覆盖范围$[a,b]$的桶标签的最小集合，假设此桶标签的最小集合为$\{T_2,T_3\}$，基站将$\{t,\langle T_2,T_3\rangle\}$代替具体范围信息发送给存储节点，存储节点收到$\{t,\langle T_2,T_3\rangle\}$后，将时间段 $t$ 内标签为 $T_2$ 和 $T_3$ 的加密数据返回给基站. 基站收到加密数据后进行解密，计算精确查询结果，将查询结果返回给用户，同时检验是否每个节点对应的 $T_2$ 和 $T_3$ 都有加密数据或验证编码上传至基站，如不是则判断为查询结果不完整，存储节点可能被俘获且删除了数据. 因为使用传感器节点和基站共享密钥进行加密，存储节点不能解密，并使用桶模式对数据进行泛化处理，实现了隐私保护范围查询.

Shi 等人的方案[25]同样采用桶模式完成隐私保护范围查询，其主要贡献为提出时空交叉验证方法(spatiotemporal crosscheck approach)对查询结果进行验证. 时空交叉验证方法由空间交叉验证技术和时间交叉验证技术组成. 空间交叉验证技术主要思路为：传感器节点生成表明采集数据在桶中分布情况的数据索引(data index)，在数据提交前广播该

数据索引,其它节点收到后将数据索引嵌入本节点上传数据中,基站利用数据索引进行查询结果完整性验证.时间交叉验证技术在上传数据中嵌入本节点表明不同时间段数据分布的数据索引,使用此数据索引可以对存储节点删除某时间段内全部数据的行为进行检测.Zhang 等人的方案[26]对时空交叉验证方法[25]进行了扩展,使用概率技术平衡隐私保护和能量消耗,并将该方法应用于多维数据范围查询.

### 6.1.2 基于前缀成员验证技术协议

SafeQ(secure and efficient query)[27]使用前缀成员验证技术(prefix membership verification technique)[38-39]在编码数据上实现范围查询而不泄漏数据值,使用加密数据链技术完成完整性验证.对 SafeQ 基本过程描述如下:设传感器节点 $s_i$ 在时间段 $t$ 采集 $n$ 项感知数据 $d_1,d_2,\cdots,d_n$,使用传感器节点 $s_i$ 与基站共享的密钥 $k_i$ 进行加密,得到加密数据项 $(d_1)_{k_i},(d_2)_{k_i},\cdots,(d_n)_{k_i}$.使用前缀成员验证技术构造具有特定属性的函数 $\mathcal{H}$、$\mathcal{G}$ 和 $\mathcal{F}$.传感器节点 $s_i$ 将 $(d_1)_{k_i},(d_2)_{k_i},\cdots,(d_n)_{k_i}$ 和 $\mathcal{H}(d_1,d_2,\cdots,d_n)$ 传送至存储节点进行存储.基站将用户提交的范围查询 $\{t,[a,b]\}$ 转换为 $\{t,\mathcal{G}([a,b])\}$ 并传送至存储节点.存储节点根据函数 $\mathcal{F}(j,\mathcal{H}(d_1,d_2,\cdots,d_n),\mathcal{G}([a,b]))$ 是否为真判断 $d_j(1\leq j\leq n)$ 是否属于 $[a,b]$,从而确定 $(d_1)_{k_i},(d_2)_{k_i},\cdots,(d_n)_{k_i}$ 中符合查询的加密数据项并将它们传送至基站,基站解密后返回用户.函数 $\mathcal{H}$、$\mathcal{G}$ 和 $\mathcal{F}$ 的以下属性保证了隐私保护范围查询的实现:(1)当且仅当 $\mathcal{F}(j,\mathcal{H}(d_1,d_2,\cdots,d_n),\mathcal{G}([a,b]))$ 为真时,有 $d_j\in[a,b](1\leq j\leq n)$;(2)存储节点不能由 $\mathcal{H}(d_1,d_2,\cdots,d_n)$ 和 $(d_j)_{k_i}$ 计算出 $d_j$;(3)存储节点不能由 $\mathcal{G}([a,b])$ 计算出 $[a,b]$.

前缀成员验证技术的关键思路就是将判断数据项是否属于某区间转化为比较数据项是否相等.基本步骤为:对数据项 $d$ 构造其前缀家族(Prefix family construction)并进行前缀数值化(Prefix numericalization)形成集合 $V$,对区间 $[a,b]$ 进行前缀转化(Prefix conversion)、前缀数值化形成集合 $U$,当且仅当 $V\cap U\neq\varnothing$ 时有 $d\in[a,b]$.图 6 演示了 SafeQ 使用前缀成员验证技术判断 $5\in[3,7]$ 的基本步骤.

SafeQ 使用加密数据链技术进行完整性验证.加密数据链技术通过对有序数据集填充冗余数据,分段加密形成数据项加密链.数据项加密链中间任何项被删除,解密后都将被检测出.以有序数据集 $\{1,3,5,7\}$ 为示例,使用密钥 $k_{i,t}$ 加密后形成的加密数据链为 $\{(\mathrm{Min}|1)_{k_{i,t}},(1|3)_{k_{i,t}},(3|5)_{k_{i,t}},(5|7)_{k_{i,t}},(7|\mathrm{Max})_{k_{i,t}}\}$.其中 Min 和 Max 分别为公共极小值和公共极大值.

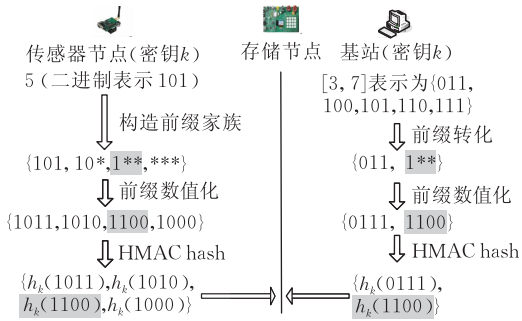


图 6 SafeQ 使用前缀成员验证技术演示

## 6.2 隐私保护 Top-$k$ 查询

SafeTQ(safe Top-$k$ query)[28]采用加随机数扰动和安全比较等技术在两层传感器网络中完成隐私保护精确 Top-$k$ 查询,使用两种完整性验证模式使基站能够检测和拒绝不正确或不完整查询响应.SafeTQ 基于与图 1 相似的网络模型,传感器网络被划分为若干个单元(cell),单元划分时应保证每个单元区域至少存在 2 个高资源节点,选择单元内一个高资源节点作为单元头节点,另外指定一个高资源节点作为辅助计算节点.SafeTQ 基于的查询模型为

$$\mathcal{Q}_t=(\mathrm{cell}=C)\wedge(\mathrm{query\ region}=G)\wedge(\mathrm{epoch}=t)\wedge(\mathrm{num}=k),$$

其中,$C$ 为查询单元,$G$ 为查询区域,$t$ 为查询周期,$k$ 为查询需要返回的最大(或最小)的数据项数.SafeTQ 仅描述查询区域在同一单元内的情况,对于查询区域为多个单元的情况能够进行分解处理.

SafeTQ 基本步骤为

(1)传感器节点 $s_i\in G$ 接到 Top-$k$ 查询 $\mathcal{Q}_t$ 后产生随机数 $r_i$,将周期 $t$ 内采集的数据集 $V_{s_i}$ 中前 $k$ 个数据分别与随机数 $r_i$ 求和,将求和结果传送到单元头节点,同时将该随机数 $r_i$ 传送到辅助计算节点.

(2)由单元头节点与辅助计算节点通过安全比较,计算出周期 $t$ 内查询区域 $G$ 中采集的全体数据集中第 $k$ 位数据值 $v_{k\mathrm{th}}$.

(3)单元头节点将 $v_{k\mathrm{th}}$ 作为阈值下发至查询区域内所有传感器节点.

(4)传感器节点 $s_i\in G$ 将 $V_{s_i}$ 中每个数据与 $v_{k\mathrm{th}}$ 比较,大于等于 $v_{k\mathrm{th}}$ 的数据即为 Top-$k$ 数据,这些 Top-$k$ 数据构成传感器节点 $s_i$ 查询响应结果集 $\mathcal{R}_{s_i}$,根据所选用的完整性验证模式进行相应处理后,使用传感器节点与基站共享密钥进行加密并传送到单元头节点,由单元头节点上传至基站.

(5) 基站解密并确定查询$\mathcal{Q}_t$结果集$\mathcal{R}_t=\bigcup_{s_i\in G}\mathcal{R}_{s_i}$(即 Top-$k$ 数据集)，同时进行完整性验证. 图 7 演示了 SafeTQ 基本步骤.

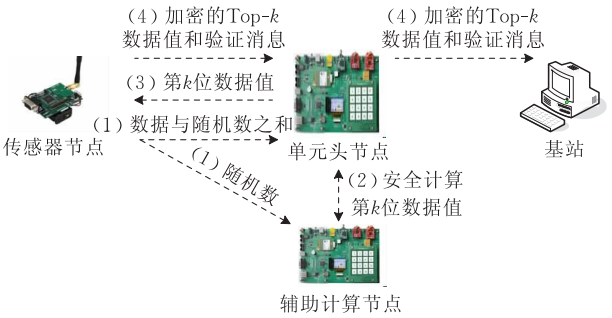


图 7　SafeTQ 基本步骤演示

SafeTQ 中两种完整性验证模式分别为数据项加密链验证模式和概率空间邻居验证模式. 数据项加密链验证模式由加密数据链技术和验证消息共同完成 Top-$k$ 查询完整性验证，使用加密数据链技术检测对$\mathcal{R}_{s_i}$的部分删除攻击，使用非 Top-$k$ 节点发送验证消息策略检测对$\mathcal{R}_{s_i}$的整体删除攻击. 概率空间邻居验证模式通过 Top-$k$ 节点的空间邻居节点以给定概率发送验证消息来完成查询结果完整性验证.

### 6.3　隐私保护基于类型查询

ElliPS(Elliptic curve based Privacy Scheme)[29]为两层传感器网络中基于类型查询(type-based queries)提供隐私保护模式. 在基于类型查询中传感器节点采集的数据属于特定类型，假设传感器节点$u$采集数据$d_u$属于类型$t_u$，基站需要查询类型为$T$的感知数据时，下发基于类型查询$\mathcal{Q}_T=\{T\}$，所有类型$t_u\in T$的感知数据$d_u$为符合查询条件数据. ElliPS 基本思路为：保证能够在存储节点完成基于类型查询的情况下，使用基于椭圆曲线多项式转换技术隐藏感知数据的类型和查询内容，同时在传感器节点和存储节点之间引入匿名处理节点(anonymizer nodes)，来应对被俘获的传感器节点和存储节点进行共谋攻击. 假设传感器节点$u$与基站共享密钥和多项式函数分别为$k_u$和$p_u(x)$，ElliPS 处理查询过程如图 8 所演示.

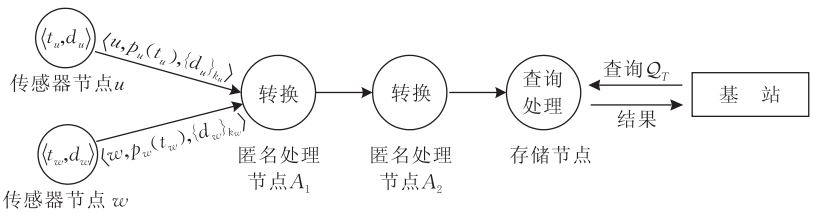


图 8　ElliPS 处理查询过程演示

### 6.4　性能分析与比较

基于桶模式协议将数据映射到桶中进行泛化处理，使用桶标签匹配实现查询，这将使得被俘获的存储节点能够对感知数据和查询范围进行合理的估计，使用前缀成员验证技术时进行这种估计比较困难[27]，所以相比于基于前缀成员验证协议，基于桶模式协议隐私保护度较弱.

在支持查询类型方面，本节所述协议为分别针对范围查询、Top-$k$ 查询和基于类型查询进行设计的，不适合直接应用于其它查询类型；在应对攻击类型方面，本节所述隐私保护范围查询协议和 SafeTQ 能够应对高资源节点被俘获发起的窃取信息攻击和篡改信息攻击，均具有隐私保护和完整性验证机制，ElliPS 在存储节点和传感器节点被俘获的情况下，能够保护感知数据的值和类型以及查询内容，但没有提供完整性验证机制；在传感器节点被俘获时，由于本节所述隐私保护范围查询协议和 SafeTQ 采用传感器节点与基站共享密钥的加密机制，传感器节点不容易获取其它节点信息；在应对共谋攻击方面，ElliPS 通过引入匿名处理节点，能够较好地应对被俘获传感器节点和存储节点发起的共谋攻击；在隐私保护度方面，SafeQ 比使用桶模式的范围查询协议具有更高的隐私保护度，SafeTQ 采用数据扰动、安全比较和端到端加密机制使其具有较强的隐私保护度；在算法性能方面，处理多维数据时随着维数的增加，Sheng & Li 的方案的能量消耗和空间代价呈指数增长，SafeQ 则呈线性增长[27]，Sheng & Li 的方案需要每个没有查询响应数据的节点返回验证信息，更适用于监测型应用，但在事件驱动型网络中可能产生不必要的高通信量[25]，Shi 等人的方案和 Zhang 等人的方案在查询响应数据中嵌入验证信息，在性能上更适应于事件驱动型网络，SafeQ 使用前缀编码技术需要较高的计算代价，SafeTQ 利用高资源节点完成安全计算第 $k$ 位数据值算法，增强了协议有效性和可行性；在结果精确度方面，本节所述协议都能得到精确查询结果. 表 3 对隐私保护数据查询代表性协议在性能方面的主要优缺点进行了总结.

表 3 隐私保护数据查询协议对比

| 代表协议 | 任务 | 技术 | 主要优点 | 主要缺点 |
|---|---|---|---|---|
| Sheng & Li 的方案[24] | 范围查询 | 桶模式 | 能够应对窃取信息和篡改信息攻击 | 隐私保护度较弱 |
| Shi 等人的方案[25] | | | 同时适用于事件驱动型网络 | 通信代价较高 |
| Zhang 等人的方案[26] | | | 适用于多维范围查询 | 概率值影响完整性验证性能 |
| SafeQ[27] | | 前缀成员验证技术 | 隐私保护度较强 | 计算代价较高 |
| SafeTQ[28] | Top-$k$ 查询 | 扰动和安全比较技术 | 隐私保护度较强 | 连续查询时通信代价较高 |
| ElliPS[29] | 基于类型查询 | 椭圆曲线多项式转换技术 | 应对共谋攻击能力强 | 没有考虑篡改信息攻击 |

## 7 隐私访问控制

在单拥有者多用户(single-owner multi-user)类型的传感器网络中,需要实施能够保护用户隐私信息的访问控制.一方面,为了数据安全或经济因素,网络拥有者需要实施严格的访问控制,只有注册并付费的用户才能访问传感器节点的感知数据;另一方面,用户希望对网络拥有者、传感器节点和其它用户隐藏其身份、访问模式等敏感信息,例如石油公司访问海洋中部署的多用户传感器网络时,希望对可能是竞争对手的其它用户隐藏其感兴趣的访问区域等信息[30].DP$^2$AC 和 Priccess 分别使用了盲签名(blind signature)技术和环签名(ring signature)技术实现了传感器网络中隐私访问控制.

### 7.1 基于盲签名协议

DP$^2$AC(Distributed Privacy-Preserving Access Control)[30]的基本过程为:用户向网络拥有者购买令牌(token),然后持令牌访问传感器节点,传感器节点验证令牌后将查询结果返回给用户.因为 DP$^2$AC 中令牌产生过程使用了盲签名技术确保了令牌是公开可验证的,并且网络拥有者和传感器节点都不能将令牌和用户身份相联系,实现了用户身份的匿名化.DP$^2$AC 包括 3 个阶段:初始化阶段、令牌购买阶段和令牌消费阶段.DP$^2$AC 使用基于 RSA 的盲签名方案[40].为表述方便设用户 U 向网络拥有者 O 购买和消费令牌.

(1)初始化阶段.网络拥有者 O 随机生成两个不相同的素数 $p$ 和 $q$,令 $n=p\times q$,$\varphi=(p-1)(q-1)$,随机选择 $e(1<e<\varphi)$,计算 $d$ 使其满足 $ed=1 \bmod \varphi$ 且 $1<d<\varphi$,得到 O 的公钥$\langle n,e\rangle$和私钥 $d$.DP$^2$AC 中公钥$\langle n,e\rangle$用来验证 O 的签名.O 将$\langle p,q,d\rangle$作为秘密信息保存,同时向用户和所有传感器节点公布公钥$\langle n,e\rangle$.

(2)令牌购买(签名)阶段.步骤 1(消息盲化):用户 U 生成随机数 $k$ 使其满足 $gcd(n,k)=1$ 且 $1<k<n-1$,U 将 $k$ 作为秘密信息保存,$k$ 为盲化因子.U 在特定范围内随机生成 $m$,计算 $m^*=mk^e$,并将 $m^*$ 传送给 O;步骤 2(签名):O 确认 U 付费后,计算 $\sigma_m^*=(m^*)^d \bmod n$,并将 $\sigma_m^*$ 发送至 U;步骤 3(去盲):U 计算$\sigma_m=k^{-1}\sigma_m^* \bmod n$,有$\sigma_m=k^{-1}\sigma_m^*=k^{-1}m^dk^{ed}=m^d \bmod n$,$\sigma_m$为 O 对 $m$ 的 RSA 签名.U 将$\langle m,\sigma_m\rangle$作为令牌.

(3)令牌消费(验证)阶段.U 持令牌$\langle m,\sigma_m\rangle$访问传感器节点 $s_i$,$s_i$ 使用公钥$\langle n,e\rangle$验证 $m$ 与$(\sigma_m)^e \bmod n$是否相等,如相等则通过验证返回 U 所需数据,因为有$(\sigma_m)^e=m^{ed}=m \bmod n$.图 9 演示了 DP$^2$AC 的基本过程.


图 9 DP$^2$AC 基本过程演示

同时,DP$^2$AC 提出分布式令牌重用检测技术 DTRD(Distributed Token-Reuse Detection),防止恶意用户重用令牌.

### 7.2 基于环签名协议

Priccess(Privacy-preserving access control)[31]使用环签名技术(ring signature)[41]实现传感器网络隐私访问控制.网络访问用户需要进行注册,网络拥有者按访问权限对注册用户进行分组,对同一组用户授予相同的访问权限,用户使用环签名技术签署查询命令并下发至传感器节点,传感器节点验证签名,确认具有相应权限后返回查询结果.环签名技术使签名验证者仅能获知签名者所属的分组,而不能获知签名者具体的身份信息,实现了用户身份的匿名并能够实施访问控制.

在传感器网络用户信息保护方面,文献[42]使用查询区域转化技术对用户查询区域进行隐藏,将查询目标区域映射到区域集合中,实现目标区域与其它区域不可分.

### 7.3 性能分析与比较

在支持查询类型方面,DP$^2$AC 和 Priccess 实现了传感器网络隐私访问控制,其中 DP$^2$AC 为该方面的首份研究成果[30];在保护的隐私内容方面,

$DP^2AC$能够在网络拥有者和传感器节点不能获知用户身份的情况下实现访问控制，Priccess 既能够实现用户身份匿名和访问控制，又能使传感器节点进行查询命令验证，能够检测出攻击者对查询命令的篡改攻击；在隐私保护度方面，$DP^2AC$使用盲签名技术，要求用户将使用盲因子计算后的信息送至网络拥有者进行签名，因为盲因子为用户私有所以其隐私保护度较高，Priccess 对注册用户的分组使用环签名技术实现了用户身份的 $k$ 匿名（$k$ 为分组中成员个数），但是攻击者能够以一定的概率（不高于 $1/k$）确定用户身份，所以 Priccess 隐私保护度比$DP^2AC$弱；在性能方面，$DP^2AC$中令牌重用检测需要较高通信量；在结果精确度方面，$DP^2AC$和 Priccess 实施隐私访问控制不影响查询结果的精确度. 表 4 对隐私访问控制代表协议在性能方面的主要优缺点进行了总结.

**表 4　隐私访问控制协议对比**

| 代表协议 | 任务 | 技术 | 主要优点 | 主要缺点 |
|---|---|---|---|---|
| $DP^2AC$[30] | 访问控制 | 盲签名 | 隐私保护度较强 | 通信代价较高 |
| Priccess[31] | | 环签名 | 提供细粒度访问控制；提供查询命令验证 | 隐私保护度较弱 |

## 8　未来工作展望

由于传感器网络数据隐私保护技术涉及多学科且发展时间较短，目前还属于新兴研究领域，很多挑战性问题有待进一步研究：

（1）复杂查询中隐私保护技术

数据查询是传感器网络的主要服务，传感器网络复杂查询中隐私保护技术的研究具有重要意义. 目前研究成果主要针对数据聚集和范围查询中隐私保护问题进行研究，而对常用重要的复杂查询类型如 TOP-$K$ 查询、JION 查询、KNN 查询、SKYLINE 查询等中隐私保护机制的研究还很少. 同时复杂查询可分为快照查询和连续查询、精确查询和近似查询等查询方式. 不同的复杂查询类型和方式往往表现出不同的特征，由于特征和实现流程不同，其可能泄漏隐私信息的内容和环节也不尽相同，而隐私保护协议往往针对特定查询有效，所以需要对不同的复杂查询类型和方式中隐私保护技术进行有针对性的研究. 所以复杂查询中隐私保护技术还有很大的研究空间.

（2）时空相关性数据隐私保护技术

传感器网络常用于环境、医疗等领域的监测，其采集的数据往往具有时空相关性. 目前传感器网络数据管理方面有部分研究成果利用感知数据时空相关性设计相应算法，以减少通信量和时间延迟. 如何将这些研究成果和隐私保护技术结合，既利用时空相关性优化聚集和查询协议性能又针对时空相关性进行数据隐私保护是有意义的研究内容. 更重要的是，攻击者能够利用感知数据时空相关性进行推断攻击. 随着时间序列模型、概率图等现代关联推断技术的迅速发展与成熟，推断攻击已成为一种新的重要攻击类型，给网络造成严重威胁. 研究传感器网络中能够应对利用数据时空相关性进行推断攻击的协议是开放性研究问题，目前此方向研究成果基本还是空白.

（3）新型网络中数据隐私保护技术

随着传感器网络技术的发展和应用领域的拓展，出现了多基站传感网、车载传感网等新型网络. 在这些新型网络中，数据存储、路由选择、聚集查询等数据管理模式将发生重大变化，将出现新的隐私保护内容和隐私泄漏渠道，攻击者可能采用新的攻击方式进行攻击. 多基站传感网中，攻击者伪装或俘获基站的可能性增大，而目前的大部分研究是基于基站可信的前提下进行的；多基站传感网可能拥有更多用户类型，用户身份和查询模式等敏感信息的保护将变得更加重要，目前关于传感器网络中用户隐私信息保护的研究才刚起步. 在车载传感网中，除数据管理模式变化外，查询的实时性显得更加重要，隐私保护数据聚集和数据查询协议应能适应这些变化以满足应用需求.

（4）隐私保护技术和传感器网络技术进一步结合

传感器网络数据隐私保护技术需要多学科的交叉和结合，主要涉及隐私保护和传感器网络数据管理. 目前，面向数据的隐私保护技术研究成果主要集中在数据挖掘和数据发布两个领域，这些研究成果为传感器网络数据隐私保护技术提供重要支撑，但传感器网络为资源受限的分布式自组织多跳网络，这些技术往往不能直接应用于传感器网络. 同时传感器网络数据管理技术的研究成果主要集中在以节省能量为目标的性能优化方面，最近几年才开始关注数据隐私泄漏问题，传感器网络数据隐私保护技术还属于新兴研究领域，还有很多需要研究的方面. 所以将隐私保护技术和传感器网络技术进一步结

合,根据传感器网络特点,在数据管理更多环节和方面进行优化,设计出性能良好的隐私保护协议,是该领域的研究方向.

(5)隐私保护度、结果精确度、能量消耗和时间延迟之间优化平衡

资源受限是传感器网络重要的特征,能量消耗量往往直接影响传感器网络的寿命.隐私保护算法往往需要较大的通信量和计算量,对能量消耗较大.传感器网络数据隐私保护协议优化目标为保证结果精确度的前提下提高隐私保护度,同时减少能量消耗和时间延迟.但是,这四者又是相互影响的,需要优化平衡.现有研究成果提出了许多优化平衡策略,但还远远不能满足传感器网络的应用需求,需要进一步研究.

**致 谢** 本文审稿专家和编辑老师提出了许多宝贵意见和建议,在此表示感谢!

## 参 考 文 献

[1] Zhou Shui-Geng, Li Feng, Tao Yu-Fei, Xiao Xiao-Kui. Privacy preservation in database applications: A survey. Chinese Journal of Computers, 2009, 32(5): 847-861(in Chinese)
(周水庚,李丰,陶宇飞,肖小奎. 面向数据库应用的隐私保护研究综述. 计算机学报, 2009, 32(5): 847-861)

[2] Agrawal R, Sriknat R. Privacy-preserving data mining//Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD). Dallas, Texas, USA, 2000: 439-450

[3] Huang Zheng-Li, Du Wen-Liang, Chen Biao. Deriving private information from randomized data//Proceedings of the ACM SIGMOD International Conference on Management of Data(SIGMOD). Baltimore, USA, 2005: 37-48

[4] Clifton C, Kantarcioglou M, Lin X, Zhu M Y. Tools for Privacy-Preserving distributed data mining. ACM SIGMOD Explorations, 2002, 4(2): 28-34

[5] Malin B, Airoldi E, Edoho-Eket S, Li Y. Configurable security protocols for multi-party data analysis with malicious participants//Proceedings of the 21st International Conference on Data Engineering (ICDE). Tokyo, Japan, 2005: 533-544

[6] Sweeney L. Achieving $k$-anonymity privacy protection using generalization and suppression. Intemational Journal on Uncertainly, Fuzziness and Knowledge-Based Systems, 2002, 10(5): 571-588

[7] Li N, Li T. $t$-closeness: Privacy beyond $k$-anonymity and $l$-diversity//Proceedings of the 23rd International Conference on Data Engineering(ICDE). Istanbul, Turkey, 2007: 106-115

[8] Shenker S, Ratnasamy S, Karp B et al. Data-centric storage in sensornets. ACM SIGCOMM Computer Communication Review, 2003, 33(1): 137-142

[9] Younis O, Fahmy S. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Transactions on Mobile Computing, 2004, 3(4): 366-379

[10] Wu Min-Ji, Xu Jian-Liang, Tang Xue-Yan, Lee Wang-Chien. Top-$k$ monitoring in wireless sensor networks. IEEE Transactions on Knowledge and Data Engineering(TKDE), 2007, 19(6): 962-976

[11] Yao Yu-Xia, Tang Xue-Yan, Lim Ee-Peng. Localized monitoring of KNN queries in wireless sensor networks. The VLDB Journal, 2008, 18(1): 99-117

[12] Oana Jurca, Sebastian Michsel, Alexandre Herrmann, Karl Aberer. Continuous query evaluation over distributed sensor networks//Proceedings of the 26th IEEE International Conference on Data Engineering (ICDE). Long Beach, California, USA, 2010: 912-923

[13] He W, Liu X, Nguyen H et al. PDA: Privacy-preserving data aggregation in wireless sensor networks//Proceedings of the 26th IEEE International Conference on Computer Communications(INFOCOM). Anchorage, USA, 2007: 2045-2053

[14] Yao Jian-Bo, Wen Guang-Jun. Protecting classification privacy data aggregation in wireless sensor networks//Proceedings of the 4th International Conference on Wireless Communication, Networking and Mobile Computing (WiCOM). Dalian, China, 2008: 1-5

[15] He Wen-Bo, Hoang Nguyen, Liu Xue et al. iPDA: An integrity-protecting private data aggregation scheme for wireless sensor networks//Proceedings of the Military Communications Conference (MILCOM). San Diego, CA, USA, 2008: 1-7

[16] Yang Geng, Wang An-Qi, Chen Zheng-Yu et al. An energy-saving privacy-preserving data aggregation algorithm. Chinese Journal of Computers, 2011, 34(5): 792-800(in Chinese)
(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. 计算机学报, 2011, 34(5): 792-800)

[17] Claude Castelluccia, Einar Mykletun, Gene Tsudik. Efficient aggregation of encrypted data in wireless sensor networks//Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services(MobiQuitous). San Diego, CA, USA, 2005: 109-117

[18] Feng Taiming, Wang Chuang, Zhang Wensheng, Ruan Lu. Confidentiality protection for distributed sensor data aggregation//Proceedings of the 27th IEEE International Conference on Computer Communications(INFOCOM). Phoenix, USA, 2008: 56-60

[19] Stavros Papadopoulos, Aggelos Kiayias, Dimitris Papadias. Secure and efficient in-network processing of exact SUM queries//Proceedings of the 27th International Conference on Data Engineering(ICDE). Hannover, Germany, 2011: 517-528

[20] Girao Joao, Westhoff Dirk, Schneider Markus. CDA: Concealed data aggregation for reverse multicast traffic in wireless sensor networks//Proceedings of the IEEE International Conference on Communications(ICC). Seoul, Korea, 2005: 3044-3049

[21] Ozdemir Suat, Yang Xiao. Integrity protecting hierarchical concealed data aggregation for wireless sensor networks. Computer Networks, 2011: 1735-1746

[22] Groat M M, He W B, Forrest S. KIPDA: *k*-indistinguishable privacy-preserving data aggregation in wireless sensor networks//Proceedings of the 30th IEEE International Conference on Computer Communications(INFOCOM). Shanghai, China, 2011: 2024-2032

[23] Zhang W S, Wang C, Feng T M. GP$^2$S: Generic privacy-preserving solutions for approximate aggregation of sensor data//Proceedings of the 6th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom). Hong Kong, China, 2008: 179-184

[24] Sheng Bo, Li Qun. Verifiable privacy-preserving range query in two-tiered sensor networks//Proceedings of the 27th IEEE International Conference on Computer Communications(INFOCOM). Phoenix, USA, 2008: 46-50

[25] Shi Jing, Zhang Rui, Zhang Yan-Chao. Secure range queries in tiered sensor networks//Proceedings of the 28th IEEE International Conference on Computer Communications(INFOCOM). Rio de Janeiro, Brazil, 2009: 945-953

[26] Zhang Rui, Shi Jing, Zhang Yan-Chao. Secure multidimensional range queries in sensor networks//Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc). New Orleans, Louisiana, USA, 2009: 197-206

[27] Chen Fei, Liu Alex X. SafeQ: Secure and efficient query processing in sensor networks//Proceedings of the 29th IEEE International Conference on Computer Communications(INFOCOM). San Diego, USA, 2010: 2642-2650

[28] Fan Yong-Jian, Chen Hong. Verifiable privacy-preserving Top-*k* query protocol in two-tiered sensor networks. Chinese Journal of Computers, 2012, 35(3): 423-433(in Chinese)
(范永健，陈红. 两层传感器网络中可验证隐私保护 Top-*k* 查询协议. 计算机学报，2012，35(3)：423-433)

[29] Nalin Subramanian, Yang Ka, Zhang Wen-Sheng, Qiao Da-Ji. ElliPS: A privacy preserving scheme for sensor data storage query//Proceedings of the 28th IEEE International Conference on Computer Communications(INFOCOM). Rio de Janeiro, Brazil, 2009: 936-944

[30] Zhang Rui, Zhang Yan-Chao, Ren Kui. DP$^2$AC: Distributed privacy-preserving access control in sensor networks//Proceedings of the 28th IEEE International Conference on Computer Communications(INFOCOM). Rio de Janeiro, Brazil, 2009: 1251-1259

[31] He Dao-Jing, Bu Jia-Jun, Zhu Sen-Cun et al. Distributed privacy-preserving access control in a single-owner multi-user sensor network//Proceedings of the 30th IEEE International Conference on Computer Communications (INFOCOM). Shanghai, China, 2011: 331-335

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

[33] Menezes A J, Oorschot P C V, Vanstone S A. Handbook of Applied Cryptography. Boca Raton: CRC Press, 1996

[34] Nath S, Yu H, Chan H. Secure outsourced aggregation via one way chains//Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD). RhodeIsland, USA, 2009: 31-44

[35] Boneh D, Eu-JinGod, Nissim K. Evaluating 2-DNF formulas on cipertexts//Proceedings of the Theory of Cryptography Conference(TCC). Cambridge, USA, 2005: 321-325

[36] Hacigumus H, Iyer B R, Li C, Mehrotra S. Executing SQL over encrypted data in the data base service provider model//Proceedings of the ACM SIGMOD International Conference on Management of Data(SIGMOD). Madison, Wisconsin, USA, 2002: 216-227

[37] Hore B, Mehrotra S, Tsudik G. A privacy-preserving index for range queries//Proceedings of the 30th International Conference on Very Large Data Bases(VLDB). Toronto, Canada, 2004: 720-731

[38] Cheng J, Yang H, Wong S H, Lu S. Design and implementation of cross-domain cooperative firewall//Proceedings of the IEEE International Conference on Network Protocols(ICNP). Beijing, China, 2007: 284-293

[39] Liu A X, Chen F. Collaborative enforcement of firewall policies in virtual private networks//Proceedings of the 27th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC). Toronto, Canada, 2008: 95-104

[40] Chaum D. Security without identification: transaction systems to make big brother obsolete. Communications of the ACM, 1985, 28(10): 1030-1044

[41] Lin Xiao-Dong, Lu Rong-Xing, Zhu Hao-Jin et al. ASRPAKE: Ananonymous secure routing protocol with authenticated key exchange for wireless ad hoc networks//Proceedings of the IEEE International Conference on Communications (ICC). Glasgow, Scotland, 2007: 1247-1253

[42] Carbunar B, Yu Y, Shi L. Query privacy in wirelesss sensor networks//Proceedings of the 4th Annual IEEE Communications Society Conference Ad Hoc Communications and Networks(SECON). San Diego, USA, 2007: 203-312

**FAN Yong-Jian**, born in 1978, Ph. D. candidate, lecturer. His research interests include wireless sensor network, privacy preservation and database.

**CHEN Hong**, born in 1965, Ph. D., professor, Ph. D. supervisor. Her research interests include database, data warehouse and wireless sensor network.

**ZHANG Xiao-Ying**, born in 1987, Ph. D. candidate. Her research interests include wireless sensor network and privacy preservation.

**Background**

Wireless Sensor Networks(WSNs) have very broad application prospects including critical area surveillance, health monitoring, military tracking, etc. As WSNs are used in some applications that include sensitive information, preserving data privacy becomes an increasingly important concern. For example, a patient's heart rate, blood pressure and other vital signs are usually of critical privacy information when monitored by medical WSNs. Hence, data privacy preservation is an essential issue in WSNs and widespread deployment of these networks could be curtailed by the lack of adequate privacy preservation.

However, resource constraints and security vulnerability are inherent limitations of WSNs. First, Sensors are usually resource-limited and power-constrained so that WSNs suffer from restricted computation, communication, and power resources. Second, because of the open nature of wireless communication channels and the lack of physical protection of individual sensor nodes, the adversary may eavesdrop on the data communication and capture sensor nodes. Hence, data privacy preservation in WSNs is a challenge.

Recently more and more attentions have been paid to this area. Extensive research has been conducted to address these limitations by developing schemes that can improve resource efficiency and enhance data privacy preservation. The content of this paper mainly provides a summary for previous works and helps researchers pay more attention to the interesting issue that need to be addressed.

This research was supported by the National Natural Science Foundation of China (Nos. 61070056, 61075053) and the Major National Science and Technology Project of China (No. 2010ZX01042-001-002-002).