物联网隐私保护研究与方法综述*

钱 萍[1a,2], 吴 蒙[1b]

(1. 南京邮电大学 a. 计算机学院; b. 通信与信息工程学院, 南京 210003; 2. 江苏科技大学 计算机科学与工程学院, 江苏 镇江 212003)

摘 要: 通过建立物联网的体系结构,详细分析了体系结构中感知层和处理层所面临的隐私安全威胁,对现有的与物联网技术相关的隐私保护方法进行了系统性的综述,重点讨论了匿名化方法、加密技术和路由协议方法的基本原理与特点,并在此基础上指出了物联网隐私保护技术今后的研究方向。

关键词: 物联网; 隐私保护; 匿名化; 同态加密; 安全多方计算

中图分类号: TP393 **文献标志码**: A **文章编号**: 1001-3695(2013)01-0013-08

doi:10.3969/j.issn.1001-3695.2013.01.003

Survey on privacy preservation in IoT

QIAN Ping[1a,2], WU Meng[1b]

(1. *a. College of Computer Science, b. College of Telecommunications & Information Engineering, Nanjing University of Posts & Telecommunications, Nanjing* 210003, *China*; 2. *College of Computer Science & Technology, Jiangsu University of Science & Technology, Zhenjiang Jiangsu* 212003, *China*)

Abstract: By establishing the architecture of IoT, this paper analyzed the privacy threats faced by sensor layer and process layer of the architecture in detail. It summarized the current privacy-preserving methods associated with IoT systemically, especially for methods of anonymization, encryption and routing protocols. At the end of the paper, it pointed out future research directions on privacy-preserving for IoT.

Key words: Internet of things(IoT); privacy preservation; anonymization; homomorphic encryption; secure multi-party computation

0 引言

物联网(IoT)近几年来受到越来越多的关注,其应用非常广泛,可以应用到军事、工业、农业、电网和水网、交通、物流、节能、环保、医疗卫生和智能家居等各个领域。然而物联网在为人们提供更多应用的同时,也面临着许多安全威胁,数据与隐私安全是物联网应用过程中的挑战之一。物联网的很多应用都与人们的日常生活相关[1],其应用过程中需要收集人们的日常生活信息(如个人的旅游路线信息、购买习惯信息等),而这些信息一般都属于个人的隐私信息。因此解决好物联网应用过程中的隐私保护问题,是物联网得到广泛应用的必要条件之一。

近年来,已经有一些研究人员开始研究物联网隐私保护方法[2,3],并讨论了物联网隐私保护的法律框架[4],同时人们对相关的网络物理系统(cyber-physical systems,CPS)隐私保护[5]和普适计算的隐私保护问题[6]也进行了研究,但目前针对物联网隐私保护的系统性研究还不多。针对这一情况,本文将从物联网的层次结构出发,分层讨论物联网面临的隐私安全问题,对物联网的隐私威胁进行了分类总结;并在此基础上对现有的与物联网相关的隐私保护方法进行综述,探讨物联网隐私保护今后的研究方向。

1 物联网概述

1.1 物联网的概念及特点

物联网可以看做是通过信息传感设备,按约定的协议实现人与人、人与物、物与物全面互连的网络,其主要特征是通过射频识别(RFID)、传感器等方式获取物理世界的各种信息,结合互联网、移动通信网等网络进行信息的传送与交互,采用智能计算技术对信息进行分析处理,从而提高对物质世界的感知能力,实现智能化的决策和控制[7]。

虽然一些比较成熟的网络和技术可以为物联网技术的研究和发展提供借鉴,但是物联网中信息的采集、处理和应用技术都与这些技术有着很大的不同,主要表现在以下几个方面:

a)与无线传感器网络相比,物联网中的感知终端数量庞大且种类繁多。物联网中可以利用 RFID、二维码、传感器、内置移动通信模块等感知、捕获、测量技术随时随地地对物体进行信息采集和获取。一般来说,无线传感器网络可以看成是物联网实现数据信息采集的一种末端网络[8]。

b)与分布式计算技术相比,物联网中负责信息处理的各

收稿日期: 2012-06-06; **修回日期**: 2012-08-14 **基金项目**: 国家"973"计划资助项目(2011CB302903);江苏省研究生创新工程资助项目(CXZZ11_0404);江苏省高校自然科学研究重点资助项目(10KJA510035);南京市重大科技计划资助项目(201103003)

作者简介: 钱萍(1978-),女,江苏镇江人,讲师,博士研究生,主要研究方向为网络安全、隐私保护(qptracy@163.com);吴蒙(1963-),男,教授,博导,主要研究方向为无线通信、信息安全.

种智能计算技术需要对海量的感知数据和信息进行分析处理，并且要实现智能化的决策和控制，即实现对事物的认知以及利用已有的信息产生新的信息。

c)与互联网提供的各种应用相比，物联网技术提供的应用将覆盖更多、更广的领域，包含了工业、农业、运输、医疗军事等许多方面。特别地，物联网的很多应用与人们的日常生活密切相关，需要收集个人的数据。

1.2 物联网的体系结构

目前人们对于物联网体系结构有一些不同的描述，但内涵基本相同。一般来说，可以把物联网的体系结构分为感知层、传输层、处理层、应用层和管理与支撑四个部分，如图1所示。

a)感知层的任务是全面感知外界信息，通过各种传感器节点获取各类数据，利用传感器网络或射频阅读器等网络和设备实现数据在感知层的汇聚和传输。

b)传输层把感知层收集到的信息安全可靠地传输到信息处理层。传输层的功能主要通过网络基础设施实现，如移动通信网、卫星网、互联网等。

c)处理层的任务是对传输层传输的信息进行相应的计算与处理，需要研究智能计算、并行计算、云计算和数据挖掘(data mining)等多种关键技术。

d)应用层是对智能处理后的信息的利用，是根据用户的需求建立相应的业务模型，运行相应的应用系统。

e)管理与支撑贯穿于各个层次中，是保证物联网实现“可运行—可管理—可控制”的关键，包括测量分析、网络管理和安全保障等方面。

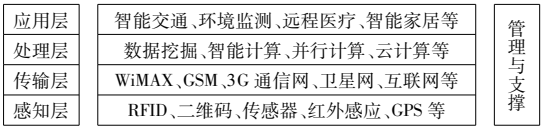

图1 物联网体系结构

2 隐私保护概述

2.1 隐私保护的概念

在研究物联网隐私保护问题之前，首先讨论隐私保护的基本概念。简单地说，隐私保护就是使个人或集体等实体不愿意被外人知道的信息得到应有的保护。隐私包含的范围很广，对于个人来说，一类重要的隐私是个人的身份信息，即利用该信息可以直接或者间接地通过连接查询追溯到某个人；对于集体来说，隐私一般是指代表一个团体各种行为的敏感信息。

2.2 隐私保护与信息安全的关系

与隐私保护密切相关的一个概念是信息安全，两者之间有一定的联系，但两者关注的重点不同。信息安全关注的主要问题是数据的机密性、完整性和可用性，而隐私保护关注的主要问题是看系统是否提供了隐私信息的匿名性。通常来讲，隐私保护是信息安全问题的一种，可以把隐私保护看成是数据机密性问题的具体体现。例如，如果数据中包含了隐私信息，则数据机密性的破坏将造成隐私信息的泄露。

本文将重点讨论跟隐私保护相关的问题，因此对于一些与信息安全交叉的问题，如数据机密性的问题，本文将不作讨论。

2.3 隐私保护方法的性能评估

通常可以从隐私性、数据准确性、延时和能量消耗这几个方面对隐私保护方法的性能进行评估。

a)隐私性是指隐私保护方法对隐私信息的保护程度。

b)数据准确性是指使用了隐私保护方法后，所能获得数据的准确性。例如在数据挖掘中，为了保护隐私信息，有时需要对原始数据进行随机化或匿名化处理后再进行挖掘，这种情况下的数据挖掘结果与直接对原始数据进行挖掘的结果相比将有所差别，即挖掘结果的准确性受到影响。

c)延时是指实现隐私保护方法时产生的延时，包括计算延时和通信延时。

d)能量消耗是指实现隐私保护方法时产生的额外能量消耗，包括数据处理和传输过程中消耗的资源。

3 物联网面临的隐私安全威胁

3.1 概述

由于物联网的很多应用需要收集个人信息，因此与现有的各种技术相比，物联网面临更严重的隐私安全威胁。与无线传感器网络相比，物联网感知终端的种类更多且数量庞大，个人的信息更容易被收集。与传统Internet相比，在传统Internet中个人可以通过在终端进行设置以保护自己的隐私，而物联网中个人无法控制自己的个人信息不被泄露；传统Internet中的隐私问题通常只与Internet用户相关，而物联网中，即使是那些没有使用物联网服务的用户，也同样存在隐私问题[1]。

一个典型的情况，考虑一个包含了多个摄像头的物联网系统[1]，该系统中摄像头获取的个人图像属于个人的隐私。在这种情况下，一个个人无法控制自己的图像不被收集，如果个人想阻止摄像头获取自己的图像，唯一的办法是不进入该区域。

考虑另一种情况，在一幢建筑中建立了一个能控制照明和采暖设备的物联网系统[1]，以实现舒适和节能的目的。该系统中放置了一些能追踪办公室内员工位置的感知设备，通过这些感知设备所采集的信息可以相应地控制开关灯或调节采暖设备。在这类系统中面临的隐私安全威胁包括：a)个人用户的位置或移动信息属于个人的私有信息，感知系统收集的员工位置或移动信息如果与具体身份相联系将侵犯到个人的隐私；b)感知系统收集的数据在传输和处理过程中面临隐私泄露的威胁。

从这两个例子中可以看出，物联网隐私保护要解决的一个主要问题是保证个人能控制自己的哪些私人信息可以被收集、由谁来收集以及在什么时间被收集[1]。

3.2 物联网隐私威胁分类

参考无线传感器网络中隐私保护的分类方法，根据隐私保护的对象来分，物联网的隐私威胁可以简单地分为两大类。

1)基于数据的隐私威胁

数据隐私问题主要是指物联网中数据采集、传输和处理等过程中的秘密信息泄露，从物联网体系结构来看，数据隐私问题主要集中在感知层和处理层，如感知层数据聚合、数据查询

和 RFID 数据传输过程中的数据隐私泄露问题,处理层中进行各种数据计算时面临的隐私泄露问题。

数据隐私往往与数据安全密不可分,因此一些数据隐私威胁可以通过数据安全的方法解决,如 2.2 节中所述,只要保证了数据的机密性就能解决隐私泄露问题,但有些数据隐私问题则只能通过隐私保护的方法解决。

2)基于位置的隐私威胁

位置隐私是物联网隐私保护的重要内容,主要指物联网中各节点的位置隐私以及物联网在提供各种位置服务时面临的位置隐私泄露问题,具体包括 RFID 阅读器位置隐私、RFID 用户位置隐私、传感器节点位置隐私以及基于位置服务中的位置隐私问题。

3.3　物联网隐私威胁分析

从前面的分析可以看出,物联网的隐私保护问题主要集中在感知层和处理层,下面将分别分析这两层所面临的隐私安全威胁。

3.3.1　物联网感知层隐私安全分析

感知层的数据一般要经过信息感知、获取、汇聚、融合等处理流程,不仅要考虑信息采集过程中的隐私保护问题,还要考虑信息传送汇聚时的隐私安全。感知网络一般由传感器网络、RFID 技术、条码和二维码等设备组成,目前研究最多的是传感器网络和 RFID 系统。

a)RFID 系统的隐私安全问题。RFID 技术的应用日益广泛,在制造、零售和物流等领域均显示出了强大的实用价值,但随之而来的是各种 RFID 的安全与隐私问题[9,10]。主要表现在以下两个方面:

(a)用户信息隐私安全。RFID 阅读器与 RFID 标签进行通信时,其通信内容包含了标签用户的个人隐私信息,当受到安全攻击时会造成用户隐私信息的泄露。无线传输方式使攻击者很容易从节点之间传输的信号中获取敏感信息,从而伪造信号。例如身份证系统中,攻击者可以通过获取节点间的信号交流来获取机密信息、用户隐私,甚至可以据此伪造身份;如果物品上的标签或读写设备(如物流、门禁系统)信号受到恶意干扰,很容易形成隐私泄露,从而造成重要物品损失。

(b)用户位置隐私安全。RFID 阅读器通过 RFID 标签可以方便地探知到标签用户的活动位置,使携带 RFID 标签的任何人在公开场合被自动跟踪,造成用户位置隐私的泄露;并且在近距离通信环境中,RFID 芯片和 RFID 阅读器之间通信时,由于 RFID 芯片使用者距离 RFID 阅读器太近,以至于阅读器的地点无法隐藏,从而引起位置隐私问题。

b)传感器网络中的隐私安全问题。传感器网络包含了数据采集、传输、处理和应用的全过程,面临着传感节点容易被攻击者物理俘获、破解、窜改甚至部分网络为敌控制等多方面的威胁,会导致用户及被监测对象的身份、行踪、私密数据等信息被暴露。由于传感器节点资源受限,以电池提供能量的传感器节点在存储、处理和传输能力上都受限制,因此需要复杂计算和资源消耗的密码体制对无线传感网络不适合,这就带来了隐私保护的挑战。

从研究内容的主体来分,无线传感器网络中的隐私问题可分为面向数据的隐私安全和面向位置的隐私安全[11,12]。无线传感器网络的中心任务在于对感知数据的采集、处理与管理,面向数据的隐私安全主要包括数据聚合隐私和数据查询隐私。定位技术是无线传感器网络中的一项关键性基础技术,其提供的位置信息在无线传感器网络中具有重要的意义,在提供监测事件或目标位置信息、路由协议、覆盖质量及其他相关研究中有着关键性的作用。然而,节点的定位信息一旦被非法滥用,也将导致严重的安全和隐私问题;并且节点位置信息在无线传感器网络中往往起到标志的作用,因此位置隐私在无线传感器网络中具有特殊而关键的地位。

3.3.2　物联网处理层隐私安全分析

物联网时代需要处理的信息是海量的,需要处理的平台也是分布式的,在分布式处理的环境中,如何保护参与计算各方的隐私信息是处理层所面临的隐私保护问题。这些处理过程包括数据查询、数据挖掘和各种计算技术等。

基于位置的服务是物联网提供的基本功能,包括定位和电子地图等技术[13]。基于位置服务中的隐私内容涉及两个方面,即位置隐私和查询隐私。位置隐私中的位置是指用户过去或现在的位置;而查询隐私是指敏感信息的查询与挖掘,即数据处理过程中的隐私保护问题。

数据挖掘是指通过对大量数据进行较为复杂的分析和建模,发现各种规律和有用的信息,其可以被广泛地用于物联网中。但与此同时,误用、滥用数据挖掘可能导致用户数据,特别是敏感信息的泄露。目前,隐私保护的数据挖掘已经成为一个专门的研究主题,数据挖掘领域的隐私保护研究最为成熟[14],很多方法可以为物联网中其他领域的隐私保护研究所借鉴。

分布式处理中要解决的隐私保护问题主要是指,当有多个实体以私有数据参与协作计算时如何保护每个实体私有数据的安全[15]。也就是说,当需要多方合作进行计算时,任何一方都只知道自己的私有数据,每一方的私有数据不会被泄露给其他参与方,且不存在可以访问任何参与方数据的中心信任方,当计算结束时,各方只能得到正确的最终结果,而不能得到他人的隐私数据。

4　物联网隐私保护方法

4.1　物联网隐私保护方法分类

对于隐私保护技术,国内外学者已经有了很多研究。目前的隐私保护技术主要集中在数据发布、数据挖掘以及无线传感网等领域,文献[15]中对数据发布和数据挖掘领域的隐私保护技术进行了总结,将其分为基于数据失真的技术、基于数据加密的技术和限制发布的技术,其中限制发布的技术主要通过数据匿名化来实现。文献[11]从数据查询隐私保护、数据聚合隐私保护和位置隐私保护等几个方面对无线传感器网络的隐私保护技术进行了综述,分别介绍了基于加密技术和路由协议技术的隐私保护方法,其中路由协议方法主要是用于位置隐私保护。在这些技术的研究基础上,结合 2.2 节划分的数据隐私和位置隐私两类物联网隐私威胁,本文将物联网隐私保护方法分为三类:

1)匿名化方法

该方法通过模糊化敏感信息来保护隐私[15],即修改或隐藏原始信息的局部或全局敏感数据。

2)加密方法

基于数据加密的保护方法中,通过密码机制实现了他方对原始数据的不可见性以及数据的无损失性,既保证了数据的机密性,又保证了数据的隐私性。加密方法中使用最多的是同态加密技术和安全多方计算(secure multi-party computation, SMC)。

同态加密最初由 Rivest 等人[16]于 1978 年提出,是一种允许直接对密文进行操作的加密变换技术,后来由 Domingo 等人作了进一步的改进,该算法的同态性保证了用户可以对敏感数据进行操作但又不泄露数据信息。

秘密同态技术是建立在代数理论之上的,其基本思想如下:

假设 E_{k1} 和 D_{k2} 分别代表加密和解密函数,明文数据是有限集合 $M = \{m_1, m_2, \cdots, m_n\}$,$\alpha$ 和 β 代表运算,若

$$\alpha(E_{k1}(m_1), E_{k1}(m_2), \cdots, E_{k1}(m_n)) = E_{k1}(\beta(m_1, m_2, \cdots, m_n)) \quad (1)$$

成立,则称函数族$(E_{k1}, D_{k2}, \alpha, \beta)$为一个秘密同态。从式(1)中可以看出,为了保护 $m_1, m_2, \cdots, m_n$ 等原始隐私数据在进行 β 运算的时候不被泄露,可以对已加密数据 $E_{k1}(m_1), E_{k1}(m_2), \cdots, E_{k1}(m_n)$ 进行 α 运算后再将结果解密,其得到的最终结果与直接对原始数据进行 β 运算得到的结果是一样的。

SMC 是指利用加密机制形成交互计算的协议,可以实现无信息泄露的分布式安全计算[15]。参与安全多方计算的各实体均以私有数据参与协作计算,当计算结束时,各方只能得到正确的最终结果,而不能得到他人的隐私数据;也就是说,两个或多个站点通过某种协议完成计算后,每一方都只知道自己的输入数据和所有数据计算后的最终结果。

3)路由协议方法

路由协议方法主要用于无线传感网中的节点位置隐私保护,无线传感网的无线传输和自组织特性使得传感器节点的位置隐私保护尤为重要。

路由协议隐私保护方法一般基于随机路由策略,即数据包的每一次传输并不都是从源节点方向向汇聚节点方向传输的[11],转发节点以一定的概率将数据包向远离汇聚节点的方向传输。同时传输路径不是固定不变的,每一个数据包的传输路径都随机产生。这样的随机路由策略使得攻击者很难获取节点的准确位置信息。

4.2 匿名化技术在物联网隐私保护中的具体应用

4.2.1 无线传感网位置隐私保护

根据节点位置的可移动性,无线传感器网络的位置隐私保护可分为固定位置隐私保护和移动位置隐私保护。针对固定节点位置的隐私保护研究较多,而移动节点位置隐私保护的研究还较少。

文献[17]中设计了一个基于匿名技术的移动位置监控系统 Tinycasper,该系统用伪装的空间位置来匿名节点的真实位置,从而保护节点的位置隐私。利用该系统,可以在监控无线传感网内移动对象的同时保护对象的位置隐私。

4.2.2 位置服务隐私保护

基于位置的服务(LBS)是物联网提供的一个重要应用,当用户向位置服务器请求位置服务(如 GPS 定位服务)时,如何保护用户的位置隐私是物联网隐私保护的一个重要内容。

利用匿名技术[18~20]可以实现对用户位置信息的保护,具体方法如下:

a)在用户和 LBS 之间采用一个可信任的匿名第三方,以匿名化用户信息;

b)当需要查询 LBS 服务器,向可信任的匿名第三方发送位置信息;

c)发送的信息不是用户的真实位置,而是一个掩饰的区域,包含了许多其他的用户。

国内学者也已经有了一些相关研究,潘晓等人对匿名的方法进行了改进,在文献[21]中提出了一种贪心匿名算法以保护查询时的用户位置隐私。这类方法的缺点是所有用户必须信任匿名的第三方,容易引起单点攻击。

4.2.3 数据查询隐私保护

数据查询是物联网提供的另一项重要服务,为了避免数据查询时的隐私泄露,可以采用数据匿名化方法,即通过将原始数据进行匿名化处理,使得数据在隐私披露风险和数据精度之间进行折中,从而兼顾数据的可用性和数据的隐私安全性,目前研究较多的是 k-匿名方法。

朱青等人[22]采用改进的 k-匿名算法,直接通过匿名化数据计算准标志符对敏感属性效用,在满足用户查询服务的同时有效地保护了数据隐私。

除了以上提到的几种应用,匿名化技术还可以用于隐私保护数据挖掘。

4.2.4 小结

匿名化技术用于数据隐私保护时,会在一定程度上造成原始数据的损失,从而影响了数据处理的准确性,并且所有经过干扰的数据均与真实的原始数据直接相关,降低了对隐私数据的保护程度;该方法用于位置隐私保护时,如 LBS 中,由于需要信任匿名的第三方,安全性不够,从而降低了隐私保护程度。

该类方法的优点在于计算简单、延时少、资源消耗较低,并且该类方法既可用于数据隐私保护,也可用于位置隐私保护。例如无线传感器网络中移动节点位置隐私保护和 LBS 的位置保护,数据处理中的数据查询和数据挖掘隐私保护等,因此在物联网隐私保护中具有较好的应用前景。

4.3 加密技术在物联网隐私保护中的具体应用

4.3.1 RFID 隐私保护

如 3.3 节中所述,RFID 主要面临阅读器位置隐私、用户信息隐私和用户位置隐私等隐私问题,下面介绍几种对应的隐私保护方法。

1)安全多方计算

针对 RFID 阅读器位置隐私,一个有效方法是使用 SMC 的临时密码组合保护并隐藏 RFID 的标志[2]。

2)基于加密机制的安全协议

对于用户的数据、位置隐私问题以及防止未授权用户访问 RFID 标签的研究,主要基于加密机制实现保护[23]。

密码机制的主要研究内容是利用各种成熟的密码方案和机制来设计与实现符合 RFID 安全需求的密码协议,安全性好,但增加了技术消耗,主要包括以下几类安全协议。

(1)基于 hash 函数的方法

a)Hash 锁协议[24]。为了避免信息泄露和被追踪,hash 锁协议使用 metaID 来代替真实的标签 ID,标签对阅读器进行认证之后再将其 ID 发送给阅读器。这种方法在一定程度上防止了非法阅读器对标签 ID 的获取,但每次传送的 metaID 保持不变,容易受到攻击。

b)随机化 hash 锁协议[25]。为了改进 hash 锁协议的不足,随机化 hash 锁协议采用基于随机数的挑战——应答机制,标签每次发给阅读器的认证信息是变化的。

c)Hash 链协议[26]。Hash 链协议是基于共享秘密的挑战——应答协议,在 hash 链协议中,要求标签使用两个不同的杂凑函数,阅读器发起认证时,标签总是发送不同的应答。

国内学者也开展了相关研究,丁振华等人[27]基于 hash 函数设计了一个介于 RFID 标签和后端服务器之间的安全认证协议 HSAP,以解决假冒攻击、重传攻击、追踪和去同步化等安全问题。

Hash 函数计算量小、资源损耗低,且 hash 函数的伪随机性和单向性保证了 RFID 标签的安全性,能有效防止标签信息的泄露和追踪。但在 hash 链中认证时,服务器端的负载会随着标签数目的增加而成比例地增长。

(2)重加密方法

重加密方案基于公钥加密体制实现重加密(即对已加密的信息进行周期性再加密),标签可以在用户请求下通过第三方数据加密装置定期地对标签数据进行重写[28]。

该方法中,由于标签和阅读器间传递的加密 ID 信息变化很快,使得标签电子编码信息很难被盗取,非法跟踪也很难实现,从而获得较高的隐私性和灵活性;但其使用公钥加密机制,运算量大、资源需求较多。

(3)匿名 ID 方法

文献[29]中提出了匿名 ID 方法以保护 RFID 用户的数据和位置隐私。该方法中标签存储的是匿名 ID,具体方法如下:

a)当标签对阅读器进行响应时,发送匿名 ID 给阅读器;

b)阅读器把收到的匿名 ID 转发给后台服务器,由服务器进行解密;

c)服务器把解密后的 ID 发送给阅读器。

该方案通过第三方数据加密装置生成匿名标签 ID,其实施前提是阅读器与后台服务器的通信建立在可信通道上,隐私侵犯者即使在消息传递过程中截获标签信息也不能获得标签的真实 ID。

该方法通过加密标签 ID 防止标签隐私信息的泄露,加密装置可以采用添加随机数等方法,资源消耗低、灵活性好。但为了防止用户的位置信息被追踪,需要定期更新标签中已加密的 ID,如果更新时间间隔太长,则隐私保护性能将大大降低。

3)其他方法

针对现有方法的不足,近年来国内一些学者还提出了其他的 RFID 隐私保护方法,邓淼磊等人[30]提出了采用伪随机函数原语实现的基于通用可组合安全模型的低成本 RFID 匿名认证协议,张辉等人[31]提出了基于部分 ID、CRC 校验以及 ID 动态更新的 RFID 相互认证协议,这些方法都能够有效地解决 RFID 安全隐私问题。周永彬等人[23]对现有 RFID 安全协议进行了分析,提出了适用于 RFID 系统环境的协议模型,对于设计和分析安全的 RFID 协议具有重要的现实和理论意义。

4.3.2　无线传感网数据隐私保护

基于加密技术的无线传感器网络数据隐私保护方法主要是采用同态加密技术实现端到端数据聚合隐私保护[32,33]。

在 WSN 中对数据进行端到端加密可以保证数据的隐私性,因此数据聚合隐私保护的挑战在于如何使聚合节点在不能解密数据的前提下对数据进行聚合。

文献[32]提出了数据聚合隐私保护方法 CDA,即利用同态加密方法使聚合节点可以对已加密数据进行聚合操作。文献[33]采用了基于加法的同态流加密算法,使得聚合节点可以对已加密数据进行聚合。

这一类方法的不足之处在于所有节点与基站共享相同的密钥,攻击者通过攻击任意一个传感器节点可以获得密钥并访问加密数据,且不能保证单个节点的隐私性。另外同态加密方法算法复杂度高,资源消耗较多。

4.3.3　数据挖掘隐私保护

针对分布式环境下的数据挖掘方法,一般通过同态加密技术和安全多方计算实现隐私保护[34],众多分布环境下基于隐私保护的数据挖掘应用都可以抽象为无信任第三方(trusted third party)参与的 SMC 问题。

下面根据数据挖掘的分类方法,从分类挖掘、关联规则挖掘和聚类挖掘三个方面介绍利用同态加密技术实现的 SMC 隐私保护数据挖掘算法[35]。

1)隐私保护分类挖掘算法

分类挖掘算法是数据挖掘中常用的一类方法。分类的目标就是要构造一个分类模型,从而预测未来的数据趋势。目前分类采用的方法主要有决策树、贝叶斯算法和 KNN 算法等。隐私保护分类技术的主要目的是要在数据挖掘过程中建立一个没有隐私泄露的、准确的分类模型。

国内外研究人员针对隐私保护分类挖掘已经展开了很多研究。Zhan[36]提出了基于同态加密和数字信封的合作决策树分类方法,参与的合作方不需要分享私有数据;Vaidya 等人[37]提出了垂直两方或多方合作下的 ID3 算法隐私保护方案;Yang 等人[38]提出了基于贝叶斯分类的隐私保护挖掘算法,可用于垂直分布的两方安全计算;Kumar 等人[39]提出了基于最近邻居查找的隐私保护方法,并利用同态加密技术在数据用户终端对私有数据进行加密。葛伟平等人[40]提出了基于转移概率矩阵隐私保护挖掘算法;张鹏等人[41]提出了基于数据处理和特征重构的朴素贝叶斯分类中的隐私保护方法。

2)隐私保护关联规则挖掘算法

关联规则挖掘是寻找在同一事件中出现的不同项的相关性,即找出事件中频繁发生的项或属性的所有子集以及它们之间应用的相互关联性。规则支持度和置信度是关注规则中的两个重要概念,它们分别代表了所发现规则的有用性和确定

性。规则 $A\Rightarrow B$ 在事务数据库 D 中成立,具有支持度 support,其中 support 是 D 中事务包含 $A\cup B$(即 A 和 B 两者)的百分比,它是概率 $P(A\cup B)$。规则 $A\Rightarrow B$ 在事务集 D 中具有置信度 confidence,D 中包含 A 的事务同时也包含 B 的百分比是 confidence,这是条件概率 $P(B|A)$。关联规则挖掘就是在事务数据库 D 中找出具有用户给定的最小支持度阈值(min_sup)和最小置信度阈值(min_conf)的规则。

Saleh 等人[42]给出了用于水平分布处理的隐私保护关联规则挖掘协议 P3ARM,该协议的关键思想是利用同态加密技术,在加密的条件下获得项集支持度。Zhan 等人[43]提出了在安全两方或多方计算的情况下,利用同态加密技术实现关联规则挖掘,即在需要用到多方数据进行挖掘计算时,用同态加密算法对各方数据进行加密。国内研究人员张鹏等人[44]将数据干扰和查询限制这两种隐私保护的基本策略相结合,提出了一种新的数据随机处理方法,并以此为基础给出了一种简单而又高效的频繁项集生成算法,进而实现了隐私保护的关联规则挖掘。

3)隐私保护聚类挖掘算法

聚类是一个将物理或抽象对象的集合分组组成由类似的对象组成的多个类的过程。由聚类所生成的簇是一组数据对象的集合,这些对象与同一个簇中的对象彼此相似,与其他簇中的对象相异,聚类分析就是从给定的数据集中搜索数据对象之间所存在的有价值联系。聚类的方法有很多,K-均值和 k-中心点是比较常用的聚类方法。

Jagannathan 等人[45]针对需要分布式安全多方计算的情况,将垂直分布和水平分布这两种概念一般化,提出了一种新的任意分布的概念,并在此基础上给出了一种可用于 K-均值聚类挖掘的隐私保护协议。Bunn 等人[46]提出了两方聚类挖掘中保护隐私的方法,该方法基于同态加密技术,包含了分离协议和随机值协议,分别实现两方分离和随机均匀采样。

4.3.4 其他隐私保护

4.2.2 节中提到了利用匿名化技术可以实现 LBS 隐私保护,但这类方法需要一个可信任的第三方,降低了安全性。Ghinita 等人[47]提出了基于隐私信息恢复(PIR)的隐私保护方法,该方法不需要一个可信任的第三方,通过加密技术实现对位置隐私的保护,并通过使用数据挖掘技术来优化查询过程。

羌卫中等人[48]对网格访问控制机制中网格实体的访问控制策略和证书的隐私保护进行了研究,提出了利用安全函数计算和同态加密理论来解决访问控制过程中策略和证书的隐私保护问题。

4.3.5 小结

用于隐私保护的加密机制一般都基于公钥密码体制(如同态加密技术等),其算法复杂度通常要高于其他基于共享密钥的加密技术,也高于一般的扰乱技术,计算延时长,且资源消耗较多。

加密机制的优点在于加密算法保证了数据的隐私性和准确性。因为利用同态加密技术的同态性质,可以在隐私数据加密的情况下对数据进行处理,既保证了数据的隐私性,又保证了数据处理结果的准确性。该类方法在现有的隐私保护技术中得到了广泛的应用,如无线传感器网络中端到端加密的数据聚合和隐私保护数据挖掘等。

4.4 路由协议方法在物联网隐私保护中的具体应用

路由协议方法主要用于无线传感网中的节点位置隐私保护。根据保护范围不同,节点位置隐私保护可分为本地位置隐私保护和全局位置隐私保护。

1)本地位置隐私保护

针对无线传感器网络的位置隐私保护,研究较早的是 Ozturk 和 Kamat 等人[49]提出的幻影路由协议。

幻影路由协议中包含了一个“熊猫—猎人”模型,作为恶意攻击方的猎人希望通过对 WSN 中无线传播信号的监测而逐跳追踪获取熊猫出现的位置。协议由两阶段构成,第一阶段是直接随机漫步,将报文随机漫步到网络中的一个伪源节点;第二阶段将报文从伪源节点路由到 sink 节点。

幻影路由协议基于随机路由策略保护节点的位置信息,但由于采用了洪泛技术,使得攻击者能很快地收集位置资源的信息。针对这一不足,国内有研究者提出了幻影路由的改进方法,Yao 等人[50]提出了定向随机幻影路由,与洪泛幻影路由不同的是,第二阶段中报文定向随机步直到基站,从而具有更大的安全期和更低的损耗。陈娟等人[51]提出了基于源节点有限洪泛的源位置隐私保护协议(PUSBRF 协议),该协议能够产生远离真实源节点且地理位置多样性的幻像源节点,从而提高了源位置隐私的安全性和平均安全时间。

2)全局位置隐私保护

Ren 等人[52]提出了可以保护本地和全局源位置隐私的路由方案。该方案由两种方法组成,路由到随机选择中间节点(RRIN)和网络混合环(NMR)。RRIN 保护本地源位置隐私,采用两步路由策略把信息从实际源节点路由到 sink 节点,通过一个或多个随机选择的中间节点使攻击者不能通过逐跳路由分析追踪到源节点;NMR 通过在一个网络混合环中路由可保护网络级(全局)源位置隐私。

文献[53]提出了可以对抗全局攻击的 sink 节点位置隐私保护方法——DCARPS 匿名路由协议,该协议中提出了一个新的网络拓扑发现方法,允许 sink 节点获得全局拓扑而不泄露自己的位置,sink 负责所有的路由计算,该协议的另一大特点是使用标记交换方法,传感器节点在转发包时执行简单的标记交换。

无线传感器网络中的主要资源消耗在于通信模块,而路由协议保护方法需要发送大量额外的通信量以实现隐私保护,因此通信开销大、能量消耗多,且通信延时长。目前的路由协议研究主要是集中于抵抗外部攻击,特别是外部攻击中的本地攻击,针对全局攻击、内部攻击和移动节点位置保护的研究相对较少,隐私保护程度不高。

4.5 其他物联网隐私保护方法

物联网的应用非常广泛,面临着各种不同的隐私威胁。为了保证收集到的个人隐私只能被用于支持授权的服务,文献[54]中提出了基于隐私代理系统的解决方案,该方案中代理一方面与用户联系,另一方面与服务提供者联系,从而保证了提供者只能获得必需的用户信息,并且用户可以设置代理的优

先权,设置和控制隐私代理使用的策略。

5　结束语

结合前面讨论的物联网体系结构和隐私安全威胁,总结物联网隐私保护方法如表 1 所示。表 2 则从隐私性、数据准确性、延时和能量消耗这几个方面,对三类隐私保护方法进行了进一步的对比。从表 2 中可以看出,不同的隐私保护方法各有其特点,可对应物联网中不同的隐私保护需求。对于物联网的感知部分,由于物联网所连接的很多终端设备的资源非常有限,因此要考虑使用计算和通信资源消耗较少的方法,如匿名化方法;对于物联网的数据处理部分,当对数据处理结果的准确性要求较高时,应考虑采用加密技术实现隐私保护。

表 1　物联网隐私保护方法

体系结构层次	面临的隐私威胁		隐私保护方法
应用层	访问控制隐私问题 其他物联网应用中的隐私威胁		隐私代理 同态加密技术
处理层	数据挖掘和分布式处理隐私问题 数据查询的隐私问题 基于位置服务的隐私问题		同态加密技术 安全多方计算 基于数据失真、匿名化
传输层	现有通信网中存在的隐私威胁因素 跨网络架构信息传输的隐私威胁		
感知层	RFID 系统	阅读器位置隐私 用户信息隐私 用户位置隐私	安全多方计算 hash 函数 重加密
	WSN	节点位置隐私 数据聚合查询隐私	匿名 ID 路由协议 同态加密技术

表 2　隐私保护方法分类分析

方法	典型应用	主要优点	主要缺点
匿名化技术	数据查询隐私保护 数据挖掘隐私保护 LBS 位置保护	延时少 能量消耗低	存在一定程度的数据损失 影响数据处理的准确性 隐私保护程度不高
加密技术	RFID 数据隐私保护 WSN 数据聚合隐私保护 数据挖掘隐私保护	隐私保护程度好 数据准确	计算延时长 由计算复杂度引起的能量消耗高
路由协议	WSN 位置隐私保护		通信延时长 由通信开销引起的能量消耗高 隐私保护程度不高

由于物联网隐私保护的研究才刚刚开始,仍然存在着许多问题有待进一步研究:

a)进一步完善现有的隐私保护方法以适应物联网环境的需求。物联网隐私保护研究可以在现有的一些隐私保护方法的基础上展开,如 WSN 隐私保护和数据挖掘隐私保护等,但是物联网与其体系结构层次所对应的基础系统之间还是存在许多区别。

从前面的分析可以看出,同态加密技术、匿名化和路由协议是隐私保护的三类重要方法。对于同态加密技术,需要研究如何有效地降低其算法复杂度;对于匿名化技术,要处理好隐私保护效果和处理结果准确性这两者之间的平衡;对于路由协议方法,应尽量减少额外通信,以实现通信量和隐私保护程度之间的平衡。

b)针对物联网多源异构性的隐私安全研究。物联网的多源异构性使其安全面临巨大的挑战,因此如何建立有效的多网融合的隐私保护模型是今后研究的一个重要方向。主要可以从以下几个方面展开研究:研究多源异构数据的数据隐藏方法;研究物联网关系链挖掘过程中的隐私保护方法;研究具有不同隐私保护安全级别的数据处理机制和协作计算算法。

c)基于语义模型的物联网隐私保护方法研究。物联网中存在着物的信息表示形式多样化与物的信息使用主体理解能力不足之间的矛盾,而语义标注和本体的引入将大为改善物的信息的共享使用,并可通过在物联网分层统一语义模型中扩展隐私保护语义属性,对指定的私密信息进行隐藏方式或销毁方式的信息遮掩,实现物联网信息的隐私保护。针对语义物联网的研究才刚刚起步,基于语义的物联网信息表示及隐私保护方法为物联网隐私保护研究提供了一个新的发展方向。

参考文献:

[1] ATZORI L, IERA A, MORABITO G. The Internet of things: a survey[J]. Computer Networks,2010,54(15):2787-2805.

[2] OLESHCHUK V. Internet of things and privacy preserving technologies[C]//Proc of the 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology. 2009:336-340.

[3] BROLL G,RUKZIO E,PAOLUCCI M,*et al*. Perci:pervasive service interaction with the Internet of things[J]. Internet Computing, 2009,13(6):74-81.

[4] WEBER R H. Internet of things:new security and privacy challenges [J]. Computer Law & Security Review,2010,26(1):23-30.

[5] 吴蒙,丁超,杨立君. 网络物理系统安全体系架构及关键技术的研究[J]. 南京邮电大学学报:自然科学版,2010,30(4):52-56.

[6] 魏志强, 康密军, 贾东宁,等. 普适计算隐私保护策略研究[J]. 计算机学报, 2010,33(1):128-138.

[7] 江苏省人民政府. 江苏省物联网产业发展规划纲要[R]. 2010.

[8] 孙其博,刘杰,黎彝,等. 物联网:概念、架构与关键技术研究综述[J]. 北京邮电大学学报,2010,33(3):1-9.

[9] OHKUBO M,SUZUKI K,KINOSHITA S. RFID privacy issues and technical challenges[J]. Communications of the ACM,2005,48(9):66-71.

[10] 钱萍, 吴蒙. 一种基于 RFID 函数的 RFID 安全认证方法[J]. 电信科学,2011,27(10):109-111.

[11] LI Na,ZHANG Nan,Das SAJAL K,*et al*. Privacy preservation in wireless sensor networks:a state-of-the-art survey[J]. Ad hoc Networks, 2009,7(8):1501-1514.

[12] 赵宝康. 无线传感器网络隐私保护关键技术研究[D]. 长沙:国防科学技术大学,2009.

[13] 顾晶晶, 陈松灿, 庄毅. 基于无线传感器网络拓扑结构的物联网定位模型[J]. 计算机学报,2010,33(9):1548-1556.

[14] ZHANG Nan, ZHAO Wei. Privacy-preserving data mining systems [J]. Computer,2007,40(4):52-58 .

[15] 周水庚,李丰,陶宇飞,等. 面向数据库应用的隐私保护研究综述[J]. 计算机学报,2009,32(5):847-861.

[16] RIVEST R L, ADLEMAN L, DETROUZOS M L. On data banks and privacy homomorphism[M]//Foundations of Secure Computation. New York:Academic Press,1978:169-179.

[17] CHOW C Y,MOKBEL M F,HE Tian. Tinycasper: a privacy-preserving aggregate location monitoring system in wireless sensor networks

[C]//Proc of ACM SIGMOD International Conference on Management of Data. New York:ACM Press,2008:1307-1310.

[18] CHOW C Y, MOKBEL M F, AREF W G. Casper*: query processing for location services without compromising privacy[J]. ACM Trans on Database Systems,2009,34(4):1-24,45.

[19] MEYEROWITZ J, CHOUDHURY R R. Hiding stars with fireworks: location privacy through camouflage[C]//Proc of the 15th Annual International Conference on Mobile Computing and Networking. New York:ACM Press,2009:345-356.

[20] MA C Y T, YAU D K Y, YIP NK, *et al*. Privacy vulnerability of published anonymous mobility traces[C]//Proc of the 16th Annual International Conference on Mobile Computing and Networking. New York: ACM Press,2010:185-196.

[21] 潘晓,郝兴,孟小峰. 基于位置服务中的连续查询隐私保护研究[J]. 计算机研究与发展, 2010,47(1):121-129.

[22] 朱青,赵桐,王珊. 面向查询服务的数据隐私保护算法[J]. 计算机学报,2010,33(8):1315-1323.

[23] 周永彬,冯登国. RFID 安全协议的设计与分析[J]. 计算机学报, 2006,29(4):551-589.

[24] SARMA S, WEIS S A, ENGELS D W. RFID systems security and privacy implications[C]//Proc of the 4th International Workshop on Cryptographic Hardware and Embedded Systems. Berlin:Springer-Verlag,2003:454-469.

[25] WEIS S A, SARMA S, RIVEST R, *et al*. Security and privacy aspects of low-cost radio frequency identification systems[C]//Proc of the 1st International Conference on Security in Pervasive Computing. Berlin: Springer-Verlag,2004:201-212.

[26] OHKUBO M, SUZUKI K, KINOSHITA S. Hash-chain based forward secure privacy protection scheme for low-cost RFID[C]//Proc of Symposium on Cryptography and Information Security. 2004:719-724.

[27] 丁振华,李锦涛,冯波. 基于 hash 函数的 RFID 安全认证协议研究[J]. 计算机研究与发展,2009,46(4):583-592.

[28] GOLLE P, JAKOBSSON M, JUELS A, *et al*. Universal re-encryption for mixnets[C]//Proc of Cryptographers' Track at the RSA Conference. Berlin:Springer-Verlag, 2004:163-178.

[29] KINOSHITA S, HOSHINO F, KOMURO T, *et al*. Nonidentifiable anonymous-ID scheme for RFID privacy protection[C]//Proc of Computer Security Symposium. 2003:29-31.

[30] 邓淼磊,马建峰,周利华. RFID 匿名认证协议的设计[J]. 通信学报,2009,30(7):20-26.

[31] 张辉,侯朝焕,王东辉. 一种基于部分 ID 的新型 RFID 安全隐私相互认证协议[J]. 电子与信息学报,2009,31(4):853-856.

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

[33] CASTELLUCCIA C, CHAN A C F, MYKLETUN E, *et al*. Efficient and provably secure aggregation of encrypted data in wireless sensor networks[J]. ACM Trans on Sensor Networks,2009,5(3):1-36.

[34] 李荣花, 武传坤, 张玉清. 判断集合包含关系的安全计算协议[J]. 计算机学报, 2009,32(7): 1337-1344.

[35] 钱萍,吴蒙. 同态加密隐私保护数据挖掘方法综述[J]. 计算机应用研究, 2011,28(5):1614-1617.

[36] ZHAN J. Privacy preserving collaborative data mining[D]. Ottawa: University of Ottawa, 2006.

[37] VAIDYA J, CLIFTON C, KANTARCIOGLU M, *et al*. Privacy-preserving decision trees over vertically partitioned data[J]. ACM Trans on Knowledge Discovery from Data,2008,2(3):1-14.

[38] YANG Zhi-qiang, WRIGHT R N. Privacy-preserving computation of Bayesian networks on vertically partitioned data[J]. IEEE Trans on Knowledge and Data Engineering,2006,18(9):1253-1264 .

[39] KUMAR P, SINGH M, SAXENA A. HEMIN: a cryptographic approach for private k-NN classification[C]//Proc of International Conference on Data Mining. 2008:500-505.

[40] 葛伟平,汪卫,周皓峰,等. 基于隐私保护的分类挖掘[J]. 计算机研究与发展, 2006,43(1):39-45.

[41] 张鹏, 唐世渭. 朴素贝叶斯分类中的隐私保护方法研究[J]. 计算机学报,2007,30(8):1267-1275.

[42] SALEH I, MOKHTAR A, SHOUKRY A, *et al*. P3ARM: privacy-preserving protocol for association rule mining[C]//Proc of IEEE Workshop on Information Assurance. 2006: 76-83.

[43] ZHAN J, MATWIN S, CHANG Li-wu. Privacy-preserving collaborative association rule mining[J]. Journal of Network and Computer Applications,2007,30(3):1216-1227.

[44] 张鹏, 童云海, 唐世渭,等. 一种有效的隐私保护关联规则挖掘方法[J]. 软件学报,2006,17(8):1764-1774.

[45] JAGANNATHAN G, WRIGHT R N. Privacy-preserving distributed K-means clustering over arbitrarily partitioned data[C]//Proc of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining. New York:ACM Press,2005:593-599.

[46] BUNN P, OSTROVSKY R. Secure two-party K-means clustering[C]//Proc of the 14th ACM Conference on Computer and Communications Security. New York:ACM Press,2007:486-497.

[47] GHINITA G, KALNIS P, KHOSHGOZARAN A, *et al*. Private queries in location based services: anonymizers are not necessary[C]//Proc of ACM SIGMOD International Conference on Management of Data. New York:ACM Press,2008:121-132.

[48] 羌卫中, 邹德清, 金海. 网格环境中证书和策略的隐私保护机制研究[J]. 计算机研究与发展, 2007,44(1): 11-19.

[49] KAMAT P, ZHANG Y, TRAPPE W, *et al*. Enhancing source-location privacy in sensor network routing[C]//Proc of the 25th IEEE International Conference on Distributed Computing Systems. 2005: 599-608.

[50] YAO Jian-bo, WEN Guang-jun. Preserving source-location privacy in energy-constrained wireless sensor networks[C]//Proc of the 28th International Conference on Distributed Computing Systems Workshops. 2008:412-416.

[51] 陈娟, 方滨兴, 殷丽华,等. 传感器网络中基于源节点有限洪泛的源位置隐私保护协议[J]. 计算机学报,2010,33(9):1736-1746.

[52] REN Jian, LI Yun, LI Tong-tong. Routing-based source-location privacy in wireless sensor networks[C]//Proc of IEEE International Conference on Communications. 2009:1-5.

[53] NEZHAD A A, MIRI A, MAKRAKIS D. Location privacy and anonymity preserving routing for wireless sensor networks[J]. Computer Networks,2008,52(18):3433-3452.

[54] LIOUDAKIS G V, KOUTSOLOUKAS E A, DELLAS N, *et al*. A proxy for privacy: the discreet box[C]//Proc of International Conference on Computer as a Tool. 2007:966-973.