

## 抗大数据分析的隐私保护：研究现状与进展

全伟<sup>1,2</sup>, 毛云龙<sup>1,2</sup>, 陈庆军<sup>1,2</sup>, 王彬入<sup>1,2</sup>, 张保佳<sup>1,2</sup>, 仲盛<sup>1,2</sup>

(1. 南京大学计算机科学与技术系, 江苏 南京 210023; 2. 南京大学计算机软件新技术国家重点实验室, 江苏 南京 210023)

**摘要:** 隐私保护对于大数据时代的数据分享和利用起到了至关重要的作用。传统的隐私保护方法无法有效地应对在大数据环境中可能存在的数据间的深度关联带来的隐私泄露。因此, 需要能够抗大数据分析的隐私保护方法。首先, 调研了国内外学者的相关研究工作, 并将当前和抗大数据分析的隐私保护相关的工作总结为海量高维异构数据隐私保护、地理位置隐私保护以及高可用性隐私保护发布方案 3 个方面。其次, 着重介绍了南京大学 COSEC 研究团队在这方面取得的一些研究进展。最后, 得到了抗大数据分析的隐私保护的重要性和前瞻性的结论, 并介绍了未来可能的研究方向。

**关键词:** 隐私保护; 大数据; 差分隐私; 地理位置隐私

**中图分类号:** TP309

**文献标识码:** A

**doi:** 10.11959/j.issn.2096-109x.2016.00042

## Survey of big-data-analysis-resistant privacy protection

TONG Wei<sup>1,2</sup>, MAO Yun-long<sup>1,2</sup>, CEHN Qing-jun<sup>1,2</sup>, WANG Bin-ru<sup>1,2</sup>, ZHANG Bao-jia<sup>1,2</sup>, ZHONG Sheng<sup>1,2</sup>

(1. Department of Computer Science and Technology, Nanjing University, Nanjing 210023, China;

2. National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China)

**Abstract:** Privacy protection plays a significant role in big data era for data sharing and analysis. Traditional privacy protection approaches may not apply well for big data scenarios, because there may exist deep associations which can also cause privacy leakages. Therefore, privacy protection mechanisms that can defend against big data analysis is needed. A survey on related works was conducted, and the previous works were classified into three categories: privacy protection for massive high-dimensional heterogeneous data, location privacy protection and privacy protection with enhanced utility. Then, the works in this area by COSEC group from nanjing university were introduced. At last, the privacy protection against big data analysis which was important and forward-looking was concluded, and the possible directions in this research area were introduced.

**Key words:** privacy protection, big data, differential privacy, location privacy

### 1 引言

随着整个社会的信息化和网络化, 海量的数据在这个过程中被收集、存储和利用。在如此大

规模数据的环境下, 如何保护数据的隐私尤为重要。近些年, 欧美国家已经着手通过立法等非技术手段来推动数据隐私的保护。美国奥巴马政府于 2012 年 2 月宣布推动《消费者隐私权利法案》

收稿日期: 2016-03-07; 修回日期: 2016-04-03。通信作者: 仲盛, zhongsheng@nju.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61321491, No.61425024, No.61300235, No.61402223); 江苏省双创计划基金资助项目

**Foundation Items:** The National Natural Science Foundation of China (No.61321491, No.61425024, No.61300235, No.61402223), Jiangsu Province Double Innovation Talent Program

(Consumer Privacy Bill of Rights)<sup>[1]</sup>以保护在线消费者的隐私安全。另外, 欧盟也于2012年1月公布了一份类似的法规《通用的数据保护法规》(General Data Protection Regulation)<sup>[2]</sup>的提案, 其中, 明确了对个人数据隐私保护的重要性。

随着大数据分析技术的日益发展, 在大数据挖掘和分析过程中显露的隐私问题引起了广泛的重视。仅仅依赖于法律法规以及以往传统的隐私保护方法, 并不能满足如今大数据环境下的隐私保护需求, 急需新型的能够抗大数据分析的隐私保护技术与方法。抗大数据分析的隐私保护主要是指在大数据环境下, 能够抵抗深度关联分析的动态隐私保护。具体的途径就是通过大规模数据的匿名、混淆等脱敏技术以及其动态组合, 来实现抗大数据分析的隐私保护目标。

本文首先从海量高维异构数据隐私保护、地理位置隐私保护以及高可用性隐私保护发布方案这几个方面来调研相关工作的研究现状。对于海量高维异构数据隐私保护这个方面, 本文主要调研了当前在社交网络数据和基因序列数据中的隐私保护方案。社交网络的数据是高维异构数据的一个代表, 包括了用户的社交关系、个人资料等数据种类, 其存在形式又有图数据、关系型数据以及非结构数据等。基因序列数据是超高维数据的代表, 其每条记录的维度可以达到数千万。对于地理位置隐私保护这个方面, 本文首先调研了保护轨迹数据隐私的数据发布技术, 然后介绍了当前差分隐私在基于地理位置的服务(LBS, location-based service)动态位置发布隐私保护中的应用。对于高可用性隐私保护发布方案这个方面, 本文调研了当前已有的提高数据可用性的扰动技术及匿名技术。

除了国际同行已有的一些研究成果, 国内的一些研究团队也从事了相关方面的一些研究。本文着重介绍了南京大学网络合作与安全研究中心(COSEC, network cooperation and security research center)团队在抗大数据分析的隐私保护方法的研究上取得的一些成果。首先, 介绍了在抗大数据分析的安全数据发布与取证方面的工作, 其中, 涉及群智感知中的隐私问题以及差分隐私的应用等; 其次, 介绍了关于智能移动网络中的

安全隐患与保护措施的工作, 包括认知无线电、移动多方安全计算等; 最后, 介绍了关于其他一些新兴网络安全问题的研究, 包括软件定义网络(SDN, software-defined network)中的安全研究等。

## 2 研究现状

### 2.1 海量高维异构数据隐私保护

海量高维异构数据主要包含社交网络数据和基因序列数据。

社交网络吸引了大量用户使用, 而社交网络数据包含了用户的社会关系、社交习惯和个人资料等敏感信息, 将社交网络数据交给第三方研究者挖掘, 可能带来用户隐私泄露的危险<sup>[3]</sup>。社交网络数据主要以图的形式存储, 保护社交网络数据隐私的一种常见方法是简单匿名化, 即对社交网络图中的顶点(即用户)做匿名处理。简单的匿名化虽然能给分析提供很好的可用性, 但这种简单的技术并不能完全保护用户隐私<sup>[4]</sup>。

针对简单匿名化方法, 研究者提出了多种攻击方式<sup>[5~9]</sup>。去匿名化方法主要分为2种类型<sup>[4]</sup>: 一是基于映射的方法; 二是基于猜测的方法。基于映射的方法将攻击者已有的背景知识和发布的数据进行匹配, 即把已知的用户ID映射到真实图中的顶点。Backstorm等<sup>[5]</sup>以及Hay等<sup>[8]</sup>分别提出了对于不同简单匿名化设定基于映射并利用已知结构信息的去匿名化攻击方法。Narayanan和Shmatikov<sup>[6]</sup>在此基础上进一步提出了一种基于反馈和自增强方式的去匿名化算法, 该方法适用于简单匿名化设定, 是一种具有顽健性的大规模算法。与基于映射的方法相反, 基于猜测的方法将发布的数据和攻击者已有的背景知识进行匹配, 即找到真实图中顶点对应的攻击者已知集合中的一个或多个候选用户。Narayanan和Shmatikov又提出了针对稀疏图的去匿名化方法<sup>[9]</sup>, 该方法根据用户购买的物品为用户打上标签, 通过对比已有背景知识和匿名化数据识别用户, 成功实现了去匿名化。Wondracek等<sup>[7]</sup>通过一种基于浏览器记录窃取的技术来识别用户, 利用窃取的浏览器历史判断社交网络链接是否被用户访问过, 从而猜测用户的身份。为了弥补简单匿名化的缺陷, 对抗去匿名化的攻击, 基于边的一系列删除与添加的边

修改  $k$ -匿名隐私保护技术、基于随机添加删除或交换边以概率的形式防止被识别的边随机化技术以及将顶点和边聚类为超节点以隐藏个体的基于聚类的泛化技术等一系列技术被提出。以上这些方法已经被视为简单匿名化在社交网络信息隐私保护中应用的必要补充<sup>[3]</sup>。

基因测序技术近几年得到了快速的发展,现在在以可接受的成本生成高精度的基因型已经成为可能。收集和分析基因序列数据能够为包括个性化医疗服务在内的各种应用带来帮助。基因序列数据维度可以达到数千万,它与某些疾病存在特定关联、具有身份识别能力,并且能够揭示家族关系。因此,基因序列数据是一种重要的医疗隐私数据,针对其特点的隐私保护方案已经被广泛地研究<sup>[10]</sup>。由于差分隐私作为一种需要隐私消耗较高的比较严格的隐私模型,在保护基因序列数据隐私的同时常常需要很大程度地牺牲数据的效用,所以,研究者的研究重点集中于如何在保护基因数据隐私的同时,提高数据的效用<sup>[11,12]</sup>。过去的隐私保护全基因组关联分析(GWAS, genome-wide association study)技术主要针对回答特定的问题,如一对给定的单核苷酸多态(SNP, single nucleotide polymorphism)之间的关系,这其实并不符合 GWAS 过程的实际情况。在实际的 GWAS 过程中,分析者也许不能提前预知对哪些 SNP 采用哪些数据测试。对此,美国海军研究实验室的 Johnson 和 Shmatikov<sup>[12]</sup>通过研究利用差分隐私来达到成员隐私,给出了一系列实际可行的作用在 GWAS 数据集上的数据挖掘算法,得到了一种在保证基因序列数据差分隐私的前提下更准确数据挖掘结果的方案。瑞士洛桑联邦理工学院的 Tramèr 等<sup>[11]</sup>针对这个问题,基于 Li 等<sup>[13]</sup>的成员隐私(membership privacy)框架,把差分隐私放松,考虑以一定概率分布获取背景知识的对手,有效地提高了隐私保护的基因序列数据在实际使用中的效用。

一些国内的研究团队也进行了社交网络隐私保护的研究。Liu 等<sup>[14]</sup>使用贪心算法对表示社交网络关系的无向图中的边进行扰动操作,再使用最短路径求解方法得到最优解。这样,就可以使用最少的数据扰动,得到最好的隐私保护效果。

另外, Liu 等<sup>[15]</sup>还提出了另外一种数据扰动的方式,高斯随机乘法的数据扰动方式。基本的方法就是在每两个相关联的实体与一组高斯分布进行随机数相乘,合理地选择高斯分布的参数,就能够得到一个很好的结果。当然,这样的方法并不是完美的方法,不过作者在文章中证明了这样的方法是一个非常优化的方法,同时也证明了完美的扰动方法是不可能实现的。Lv 等<sup>[16]</sup>提出了另一种社交网络隐私保护方案。他们首先设计了带有陷门属性的加密算法,假设有一个主体负责用户私钥的生成与分发,另一个主体负责属主私钥的分发,然后再通过令牌树,实现了属性撤销。这样,就避免了社交网络服务提供者(SNSP, social network service provider)与系统内部非授权用户的合谋攻击。这种方法还是比较不常见的,因为 SNSP 一般都是可信任的,如果 SNSP 是不可信任的,可以选择拒绝他的服务。

## 2.2 地理位置隐私保护

随着具有记录地理位置信息功能的移动设备的迅速普及,地理位置服务商能够收集到大量的个人位置信息,对这类信息进行不恰当地挖掘和分析将导致用户个人隐私的泄露,甚至可能导致对用户人身安全的威胁。

研究者们提供了各种保护轨迹数据隐私的技术<sup>[17~20]</sup>,这些技术主要分为3类:1) 基于假数据的轨迹隐私保护,这类技术将真实数据和假数据一同发送给服务提供者,让服务提供者对于所有收到的数据进行回复,从而使服务提供者即使存储了收到的轨迹数据也不能够区分真实数据和假数据<sup>[17]</sup>,但是,假数据的引入在增强隐私保护的同时增加了通信成本,基于此,作者还提出了一个成本减少技术,从而显著地提高了该方法的实际可用性;2) 基于抑制法的轨迹隐私保护,即有选择地发布原始数据,抑制某些数据项<sup>[19]</sup>,或者根据某个区域访问对象的多少将地图上的区域分为敏感区域和非敏感区域,一旦移动对象进入敏感区域,将抑制或推迟其位置更新,以保护其轨迹隐私<sup>[20]</sup>,然而,过多的轨迹片段被抑制将导致巨大的信息损失,因此,这类方法一般仅适用于对于隐私保护要求较高而对数据可用性要求较低的场景;3) 基于泛化法的轨迹隐私保护及轨迹  $k$ -

匿名技术<sup>[18]</sup>，这类技术将轨迹上所有的位置点泛化为相对应的匿名区域，在隐私保护和可用性上达到了一种平衡，因而在轨迹数据保护中被广泛使用。

作为近几年隐私保护领域研究者密切关注的隐私定义差分隐私，是现今地理位置隐私保护中最常用的技术。Dwork 等<sup>[21]</sup>最早给出了差分隐私的标准定义，定义由具体的数学公式给出，其思想是在要发布的数据集中加入或删除一条信息，不会显著地影响任何分析的结果，即因加入或删除一条信息所泄露的隐私被控制在极小的范围内。另外，还通过在查询结果中添加随机噪声来公布数据，使攻击者不能分辨原数据，从而在满足差分隐私定义的前提下，同时满足查询的可用性需求。下面主要介绍差分隐私在基于地理位置的服务（LBS）应用中如何保护动态发布的位置隐私。

基于地理位置的服务在社会和商业领域均广泛流行，传统的  $k$ -匿名方法已经不足以保护用户位置隐私<sup>[22]</sup>，因此，差分隐私技术成为 LBS 应用用户位置隐私保护的主要方法。近年来，研究者们对这个领域的研究取得了诸多成果<sup>[23~27]</sup>。Andrés 等<sup>[23]</sup>提出了一种适用于地理位置保护的差分隐私的变形：地理位置不可区分性，这种隐私定义在一定半径内满足保护用户位置隐私的直觉并且提供根据半径大小变化的保护级别。此外，他们还提出了 planar laplacian，一种通过向用户地理位置中加入可控随机噪声来获取地理位置不可区分性的机制。Chatzikokolakis 等<sup>[24]</sup>基于上述机制，设计了能够提高机制效用和降低隐私开销的满足地理位置不可区分性的机制。此外，Chatzikokolakis 等还针对上述机制中出现的加入独立噪声会导致隐私保护的迅速失效的问题，又提出了一种基于预测函数的高效机制<sup>[25]</sup>，在提高效用和降低隐私开销的同时，满足地理位置不可区分性。Bordenabe 等<sup>[27]</sup>提出了给定位置不可区分性程度下利用线性规划实现的最小化服务质量损失的具体方法，并且给出了一种将线性规划约束从三次降至二次而不显著影响可用性的方法。针对过去的隐私保护策略大多只考虑地理位置的空间关系而忽略时间关系的特点，文献<sup>[26]</sup>给出

了一种能够严格保证隐私的系统解决方案。文中定义了“ $\delta$  位置集”差分隐私来处理位置数据的时间关联，提出了新概念“sensitive hull”来代替传统的  $l_1$ -范数敏感度，并最终给出了这种平面各向同性机制，此机制是第一个达到差分隐私下界的机制，并且同时具有较好的可用性。

国内的一些研究团队在轨迹保护上做了一些工作<sup>[28~30]</sup>。Huo 等<sup>[28]</sup>提出了一种移动社交网络中的轨迹隐私保护方法。方法的核心是在中心的服务器上实现能有效保护移动社交网络中假名用户的轨迹隐私，从而保证攻击者无法以很大的概率得到签到中的用户的数据。这种方法的主要想法就是利用用户签到的序列建立一个签到的前缀树，然后再对这颗前缀树进行剪枝操作得到满足条件的  $k$ -匿名前缀树，接下来就是遍历这颗前缀树，就可以生成满足轨迹的  $k$ -匿名的可签到的序列。Wu 等<sup>[30]</sup>提出另外一种轨迹发布方法，这种方法对前面的聚类方法进行了改进，实现了  $k$ -匿名算法与聚类轨迹发布算法的融合，达到了更好的效果。

### 2.3 高可用性隐私保护发布方案研究

高可用性隐私保护发布方案研究在保证用户数据及其关联规则等隐私信息不被泄露的前提下，发布用户的数据以进行数据挖掘或分析，并尽量提高发布数据的可用性。数据扰乱技术是这类研究中常用的技术，它的基本原理是修改原始数据，使数据挖掘者不能从最终发布的经过修改的数据中直接获取或推测出原始数据的信息。

实现具有高可用性的数据扰乱是当前研究的主要目标，现有的工作中有多种方式实现数据扰乱。Agrawal 等<sup>[31]</sup>首先提出了基于随机数值与原始数据叠加的数据扰乱方法，该方法在随机改变原始数据使数据挖掘者不能得到准确原始数据的情况下，能够提供较高的可用性。Oliveira 等<sup>[32]</sup>基于数据清洗的思想提出了隐私保护关联规则挖掘算法，通过修改或删除数据记录的方式，减少了某些频繁项集的支持度，从而隐藏了关联规则。Saygin 等<sup>[33]</sup>基于数据阻塞（也称数据屏蔽）的思想也提出了隐私保护关联规则挖掘算法，他们通过用不确定值（问号）代替隐私属性值的方法，将规则的支持度不确定化。Samarati 和 Sweeney<sup>[34]</sup>

将数据泛化技术运用到隐私保护中,使用语义一致但不精确的替代值来替代真实数据值,在确保所发布的信息数据公开可用的前提下,隐藏公开数据记录与特定个人之间的对应联系。为了应对针对匿名化常见的链接攻击(linking attack), $k$ -匿名<sup>[35]</sup>、 $l$ -多样性<sup>[36]</sup>、 $t$ -closeness<sup>[37]</sup>等多种匿名模型被提出。然而,这些匿名策略为了保护隐私,需要以损失信息为代价,使数据挖掘分析的效果大打折扣,为了提高匿名后数据的可用性,研究者们进一步提出了个性化匿名、加权匿名等一系列匿名策略。不同于传统匿名策略对所有记录执行相同的匿名保护,这些匿名策略通过对不同数据记录进行不同程度的匿名保护,很大程度上减少了信息损失,提高了数据的可用性。除了链接攻击,Narayanan等<sup>[9]</sup>还给出了一种从多个数据源获取充足信息以去匿名化的攻击方法,给现有的匿名策略提出了新的挑战。

### 3 COSEC 团队研究进展

#### 3.1 抗大数据分析的安全数据发布与取证

在网络大数据时代与数据挖掘兴起的背景下,用户数据的安全面临多方威胁且攻击方式通常复杂,难以抵御,因此,数据的安全发布及取证成为了数据安全中的重要一环。移动群智感知在移动市场以及科研领域都引起了广泛的关注。通常情况下,一位感知任务的发布者会在群智感知市场向任务参与者发布自己的感知任务。那些具有合适设备的用户,通常为移动智能设备的持有者,便可以接受并参与这项感知任务。感知任务的内容通常涉及参与者的隐私信息,因为移动设备的感知需要使用到移动设备所配备的各种传感器。众所周知,传感器设备可以感知所处环境的数据信息,如智能手机上的卫星定位传感器、重力加速度传感器、光学传感器等,这些传感器所收集的信息不仅可以暴露用户的地理位置,还有可能暴露用户所处的环境信息,甚至有最新的研究文献表明智能手机的传感器信息可以用来识别用户的身体动作。综上所述,移动感知的发展前景是不可限量的,然而其操作过程涉及到用户的重要隐私信息,这使用户对接受这一产品产生了反感情绪,甚至阻碍移动感知的发展。南京大

学 COSEC 团队在目前主流的移动感知市场模型下,同时考虑了感知任务发布者与感知任务参与者的个人隐私信息的泄露可能,提出了一种适用于各种移动感知众包任务的交易平台<sup>[38]</sup>。感知任务发布者不会暴露自己的身份,同时感知任务的参与者也被保证不会泄露任何与真实身份有关的信息,这使交易的双方都满足匿名化的需求。并且,通过本文提出的交易平台,感知参与者获得的劳动报酬也可以匿名支付与提现,并且在产生伪造虚拟货币的时候可以有完美的取证与验证协议。最后,COSEC 团队实现了提出的感知任务交易平台,并且验证了其正确性与效率。

在网络市场中广告费用是各大运营商盈利的主要手段之一,而目前国际上流行的广告付费模式存在数据伪造及抵赖的威胁。一般来说,互联网广告付费模式是以安排广告投放的代理商为中心,即广告代理商接受有投放广告需求的经销商的投放申请及要求,代理商将根据这些要求选择投放广告的目标网站,这些目标网站就是广告投放的最终实施者,也是统计广告点击量的负责人。代理商会根据广告的点击量向申请投放广告的经销商收取费用,并向发布广告的网站支付报酬。如果代理商在广告的点击量数据上作弊,就可以向经销商收取不符合实际情况的广告费用,并且由于无法验证广告点击的账单,恶意的经销商也可以抵赖拒绝支付账单。更进一步地,广告发布者即目标投放网站,也可以伪造点击量数据,来向代理商索取更多的利益分红。COSEC 团队通过对广告市场中参与者的权利重分配,提出了一种新的广告投放模式<sup>[39]</sup>,通过匿名化技术及数字签名方法生成不可伪造且不可抵赖的广告点击数据报告,实现了广告市场中可靠平等的交易。更进一步地,COSEC 团队还提出了一种追踪谎报点击数据的取证方式,以检测确定在市场中作弊的广告发布者<sup>[39]</sup>。

在现实中,一个服务商有可能需要联合别的一些服务商发布各自积累的用户数据,以供第三方机构更好地进行数据挖掘的研究。但是,用户数据包含了大量的用户隐私信息,这些信息是非常敏感的。因此,在发布数据之前,需要对数据进行匿名化处理,从而保证任何数据使用者都无

法从发布的数据中获取到用户的隐私信息。为了达到这个目的,可能会牺牲一部分数据效用,也叫做数据可用性。在这个过程中,如果数据的效用损失太大,数据使用者在后期使用数据时就会受到较大的影响。所以,数据提供者和数据使用者需要判断发布的数据是否符合预期或者使用要求。目前,还没有很好的办法来解决这一难题。COSEC 团队基于差分隐私提出了一种在多数据源场景下验证隐私保护数据发布效用的方案<sup>[40]</sup>。并针对集合型数据和关系型数据分别设计了效用验证算法。此方案仅需要数据发布者在发布匿名化数据集的同时发布一些原始数据集的统计信息。虽然需要发布的仅仅是原始数据集的统计信息,但还是可能包含一些隐私信息,所以不能直接发布。该方案利用密码学工具对这些统计信息加密并由各个数据源验证这些信息的正确性,最终验证者据此计算数据集的效用。文中通过理论分析表明该方案在整个过程中都不会违背差分隐私属性,也不会引入新的隐私问题。最后给出了该方案的实验设计,并在 2 个真实数据集中测试了其时间效率,实验表明其足够高效。

推荐系统可以用来为用户提供定制的内容或者服务的推荐,并且推荐系统能够提升网站的收益以及用户体验。但是,推荐系统需要用到一些用户的隐私信息,才能为用户带来更好的体验。推荐系统甚至可能利用用户隐私数据谋利<sup>[41]</sup>。那么在保护用户隐私的同时构建一个推荐系统就十分必要了。有许多采用协同过滤技术的推荐系统。而矩阵分解是其中最流行最成功的,它获得了 Netflix 竞赛的奖项<sup>[42]</sup>,并且已经在许多真实的推荐系统中应用。因此,COSEC 团队设计了一种隐私保护的矩阵分解机制<sup>[43]</sup>。文中对推荐系统可信与不可信分别给出了相应的解决方案。考虑到实际的矩阵分解过程中用户可能动态加入或者离开,针对这种情况,也给出了相应的解决方案。文中在两个数据集上对所提出的方案进行了实验,实验结果表明,提出的方案足够高效,并且增加隐私保护对于矩阵分解的结果影响非常有限。

轨迹数据(或称人类移动轨迹)在一些领域中是非常有价值的数据<sup>[44,45]</sup>,比如城市规划、公共交通规划。虽然轨迹数据有这些价值,但是轨

迹数据里包含了个人的隐私信息<sup>[46,47]</sup>。COSEC 团队设计了一种基于差分隐私的通用时序轨迹数据发布机制<sup>[48]</sup>,并且从理论上证明了此机制满足  $\epsilon$ -差分隐私。文中针对提出的机制设计了相应的实验,实验的数据来源于真实的数据集,数据集里包含了 6 000 条出租车轨迹。实验结果表明,提出的机制在效用和时间效率方面都有不错的效果。

### 3.2 智能移动网络中的安全隐患与保护措施

在无线通信领域,无线频谱一直是最珍贵的资源。无线通信设备的快速发展使无线频谱的分配十分困难。通过实际的经验以及学者的研究表明,无线频谱的分配与利用存在极不合理的地方。简单来讲,有一部分频谱被分配给像国家数字电视广播一类的大客户,但是这些授权客户在一段时间内对某些频谱上的使用时间十分有限,导致了资源的极大浪费。为了解决这一问题,认知无线电的观念被提出,即通过协调无频谱授权的次要用户,在授权用户的频谱空闲时间使用该频谱资源进行无线传输。这种方法很好地解决了频谱资源的浪费问题,但是也催生了一系列新问题,例如授权用户与无授权的次要用户之间的冲突与干扰问题。为了更好地感知授权用户的使用时间,协同感知被认为是认知无线电的必要部分。同任何新技术一样,协同感知的出现也引入了很多问题,其中,很重要的一个就是用户的隐私,尤其是地理位置隐私的泄露。由于次要用户的感知报文中包含可以推算出地理位置的信息,任何监听的攻击者都可以获得次要用户的地理位置,这存在极大的安全隐患。南京大学 COSEC 团队通过引入一种理想的加密技术,对次要用户的感知报文做安全加密,并在融合中心做安全的数据融合,可以在不影响最终的感知结果的基础上有效地保护次要用户的感知报文的安全,使攻击者对于次要用户的隐私攻击无从下手。在提出防御方案后<sup>[49]</sup>,COSEC 团队还考虑了更为危险的恶意攻击模型,将原有防御方案扩展到恶意攻击模型下,使安全定义的假设最小化,从而为认知无线电中协同感知的工作清除了隐私泄露这一安全隐患。更进一步地,通过仿真实验,证明了保护方案的可行性与效率。

认知无线电中另一种有效提高效率的方法是协作中继方式。通过节点的协作中继可以提高频谱的利用率。然而,恶意的节点可以通过谎报信道及费用信息来误导其余的次要用户及授权用户。这种作弊行为会损害其余次要用户的利益,并且导致很低的系统效率。COSEC 团队利用博弈理论提出了首个不可作弊的协作中继方案<sup>[50]</sup>。通过建立次要用户间的博弈模型,COSEC 团队证明了任何自私的次要用户都无法从作弊中获得任何收益,更进一步地,还保证了认知无线电系统的公平性及安全性<sup>[50]</sup>。作为无线网络中的重要资源,无线频谱通常以固定的单位在次要用户之间进行拍卖。认知无线电的出现促进了可调节的带宽利用率。因此,次要用户间的可变的频谱拍卖也迫在眉睫。COSEC 团队在文献[51]中提出了一种安全可靠的可变频谱的拍卖架构。通过这一模型,可调节频谱的拍卖及分配成为可能。并且,在多个冲突域的情况下提出了另一个相似模型并证明了其可靠性。

近年来,有大量研究证明智能移动设备可以用来实现准确的室内定位。其中,很大一部分需要使用声纳系统作为主要手段或是辅助工具。但是,声纳系统的使用给智能移动设备的室内定位带来了新的安全隐患。作为室内定位使用的声纳系统通常需要3台智能手机或其他移动设备,如果其中存在恶意用户,提供错误的定位信息,或者在系统外部存在其他的智能手机或移动设备作为攻击者,室内定位不仅会受到影响,甚至无法得到正确结果。COSEC 团队利用成对的智能手机互相确认每两台设备之间的距离,从而确定每台智能设备无法谎报目前的确切位置。Hua 等<sup>[52]</sup>利用设备的自相关信息提出了一种顽健性很好的声纳定位系统,可以抵抗系统外攻击者对定位过程的干扰。

随着时代的发展,目前的智能手机已经配备有许多传感器,如GPS传感器、加速度传感器等。智能手机感知出色、计算和通信方便,这使智能手机可以轻松地完成一些移动感知的任务。近年来,出现了一批利用智能手机来采集感知数据的项目<sup>[53-57]</sup>。从技术上来说,感知的任务可以外包给智能手机的用户,手机用户采集感知数据并传

输给任务发布者。但是,对于某些采集的数据是包含用户隐私信息的,如可能包含物理位置信息、身体状况信息等。所以,对于用户来说,如果涉及到其隐私,那么手机用户很可能因此不参与到感知任务中。

COSEC 团队提出了针对场景“数据收集者希望周期性地采集所有用户时序数据的最小或者第 $k$ 小值”来保护用户隐私的方案<sup>[58]</sup>,并且不要求数据收集者可信。基于概率编码机制和一种异或同态加密系统,COSEC 团队针对收集最小值以及第 $k$ 小值分别设计了安全的协议。文中在半诚实模型下论证了协议的安全性。此外,还针对提出的协议进行了相应的时间效率实验,实验结果表明,所提出的协议都足够高效。

目前,很多关于移动感知的隐私保护都是针对真实数据的内容进行保护,这些工作中的协议都是通过特别设计后用于计算某个特定的聚合函数,比如收集数据的最小或者第 $k$ 小值<sup>[58]</sup>、高效安全的协议来计算所有用户的时序数据总和<sup>[59]</sup>。这样的做法在需要收集多个聚合函数时需要单独设计多个协议,非常低效。而且目前大多数现有工作中的聚合函数都是类似求和、求平均值、求最大值、求最小值等简单函数,而一些如方差、F检验、Z检验这样的非线性函数很少被提及。COSEC 团队采取了完全不同的思路来保护用户的隐私——切断数据和用户的关系<sup>[60]</sup>。根据这个思路提出了匿名数据协议。这个协议让数据收集者可以周期性地收集到所有用户数据的随机排列,但无法识别任何一个数据的来源。文中针对匿名数据协议的效率进行了相应的实验,实验结果表明,匿名数据协议足够高效。

在移动感知应用中,为了激励手机用户参与到应用中,感知任务发布者可能会根据参与者所能完成的任务给予参与者一定的报酬。每个手机用户根据自己的情况进行出价,而任务发布者则有一个自己的预算。COSEC 团队提出了一种能保证最低收益的高效防护策略拍卖机制<sup>[61]</sup>。文中证明了所设计的机制能够抵御不诚实的出价以及这个拍卖机制能保证给拍卖者的最低收益,并且给出了这个最低收益与最大收益的比值。随后,文中提出了2个高效的验证算法来验证拍卖机制的

结果，并证明了这个验证算法的正确性：这2个高效的验证算法不会泄漏参与者的隐私信息，同时能够有效地防止拍卖的结果被篡改。文中设计了大量的仿真实验来评估所提出的机制。实验结果验证了机制的效率，并且拍卖所获得的收益相当优秀。

### 3.3 新兴网络安全问题研究

无线通信安全一直是无线通信领域的热门话题，然而，传统的无线通信安全协议基本都是基于传统密码学手段，而传统密码学的安全性在于计算复杂度的假设。在硬件设备与计算能力日新月异的今天，传统密码学正面临严峻的挑战。COSEC 团队基于传统密码学的计算复杂度假设这一缺陷，提出了一种基于无线信道特征的通信安全协议。具体而言是通过提取无线信道的特征构建一种二分之一不经意传输协议，从而实现加密通信信道。从传统密码学中完全移除计算复杂性的假设，构架无线安全信道是十分困难的，但这并不意味着要构建完全独立于传统密码学的安全信道，而是使用无线信道的固有特征取代传统密码学中的计算复杂性的假设。准确地讲，使用无线信道的特征数据重新构建一种重要的密码学工具<sup>[62]</sup>，即二分之一不经意传输协议。文献[62]所构建的二分之一不经意传输不基于任何计算复杂度假设，完全依靠无线信道物理特征；所构建的不经意传输协议需要具有完备性质以满足构建上层密码协议的需求；更重要的是，COSEC 团队完全实现了文中所描述的不经意传输协议，并在现实生活中实现了2个实际应用，验证了本文提出的无线安全通信协议的正确性与可行性。

另一方面，由于 Wi-Fi 的广播工作特性，攻击者很容易在 Wi-Fi 环境中窃取用户的个人隐私。然而 COSEC 团队发现了一种全新的基于 Wi-Fi 的针对用户隐私的攻击方式。攻击者之前窃取的用户数据只是从分组中得到的网络数据，然而对于用户的身份信息却很难捕捉。COSEC 团队发现通过获取 Wi-Fi 分组与摄像头的协助，攻击者可以很快地准确确认受害人身份<sup>[63]</sup>。这种攻击的实现首先是基于用户在移动过程中对可视信号，即视频信息的影响（如用户的外形特征、位置的移动、移动的模式等）和对不可见信号（如

接收信号强度）的影响。对接收信号强度的影响可以在信号的频域上得到很好的体现。通过文献[63]提出的 EV-Linker 攻击方式，攻击者可以在短时间内匹配所窃取的信息与攻击目标的真实身份。文献[63]不仅考虑了在移动中的攻击方式，也考虑了在静止时的情况，如何确定攻击目标的身份，并且实验证明，这一新型攻击方式具有非常高的准确率。

无线通信中的传输速率是用户最为关心的一个问题，对传输速率造成影响的因素非常多，其中极为关键的一个因素就是网关的选择。然而，无线通信中网关受到多方面的影响，用户很难通过简单的选择方式找到最优的网关。对于多域间的网关选择，国际上最为流行的方式是分布式的学习方式。虽然这种方式有很好的实际效果，但是存在一个根本性的缺陷，在多个纳什均衡的情况下，这种学习算法有可能长时间处在一个非最优选择的纳什均衡的解上，并且在这些非最优解上的停留时间非常长，以至于有很大可能用户无法在学习算法收敛之前到达最优的选择解。鉴于这种情况，COSEC 团队对网关的选择问题做了系统性的调研并将存在的所有可能分为3种情况。针对在公用网络中的链接，文献[63]提出了一种可以直接计算得到最优网关的算法；针对2个存在私有链接的域的情况，文献[64]提出了一种基于密码学工具的方法得到最优网关选择；而对于存在3个及以上域的情况，文献[64]提出了一种引入扰动的算法来计算最优的网关选择。在上述的3种情况中，所提出的算法都可以在合理的时间开销下给出最优的网关选择结果，并且作为一种集中式的算法，在没有分布式的计算优势下，所给出的最优选择结果不会比分布式的学习算法得到的结果差。

在软件定义网（SDN, software defined networking）中，控制平面决定了网络的转发规则。网络的管理员可以在一台集中式的控制器上对交换机的策略进行远程配置。与传统的网络架构相比，SDN 使管理员更容易对网络配置进行更新，配置通常包括增加、修改或者删除某些交换机的策略。但是由于交换机分布在不同地方，所以，没有办法保证所有的交换机都能同时接收到管理

员的更新配置。那么在配置更新完毕前,网络中的一个分组可能会同时被新的和旧的2种策略处理,这种不一致可能会造成严重的后果,比如转发环路、数据分组丢失等<sup>[65~68]</sup>。COSEC团队提出了能保证数据分组一致性的一种高效策略更新机制<sup>[69]</sup>,并在SDN的实验环境中对提出的机制进行了测试。实验结果表明,文中的机制不论是在更新速度还是在更新过程中的吞吐量都是非常优秀的。

目前,存在很多多域网络来连接数据中心、企业网、大学校园网等。在这样的多域网络中,不同的域可能属于不同的机构。有的机构或者域在与其他域进行通信的过程中,会在意自己的域内隐私信息,因此,保护域的隐私非常重要。在SDN中,控制平面与数据平面被分离开,管理员可以方便地从控制平面采用集中式的方式来对路由策略决策。但是,这仅仅只能保证单域的路由策略最优。目前,现有的文献中跨域路由的策略有2种方法:一种是在域内采用集中式的办法,在域间采用边界网关协议,如谷歌的设计B4<sup>[70]</sup>,这种域内最优加上边界网关协议的方案无法保证得到的路由策略是全局最优的;另外一种方案则是采用类似OSPF的方案来计算全局最优的路由策略,这种方案能得到全局最优,但是需要每个域都将域内的拓扑信息、链路信息等共享,显然这样做就失去了隐私,是不合适的。COSEC团队提出了一种基于隐私保护的跨域路由策略优化方案<sup>[71]</sup>。文中首先提出了一种基于隐私保护的最短路由路径的计算;接着,在此基础上,又设计了带宽分配的协议并严格证明了所提出的方案的安全性;此外,针对所提出的方案进行了时间和通信效率的实验,实验结果表明,文中的方案在时间和通信方面的表现都足够优秀。

## 4 结束语

大量的大数据处理系统和分析方法在近几年得到了飞速的发展,而关于大数据环境下的隐私保护技术却是相对滞后的。本文认为大数据时代的核心应该是数据的分享和利用,而数据分享和利用的基础便是数据共享者的数据隐私得到保护。因此,可以说抗大数据分析的隐私保护是大数据领域进一步发展的基础,是一个非常重要并

具前瞻性的研究领域。对于该领域未来的研究方向,本文提出以下3点:1)差分隐私在各个应用领域的发展有待进一步的研究;2)如何在保证数据隐私的前提下,进一步提高隐私保护的后的数据的效用是至关重要的,这一点也可以理解为如何在隐私保护、效用以及数据的应用范围之间做很好的妥协;3)如何利用、整合已有的和未来开发的隐私保护方法,构建出实际可用的抗大数据分析的隐私保护工具集和系统。

## 参考文献:

- [1] The white house. Consumer data privacy: in a networked world[R/OL].<https://www.whitehouse.gov/sites/default/files/privacy-final.pdf>.
- [2] European Commission. Proposal on general data protection regulation[R/OL].[http://ec.europa.eu/justice/data-protection/document/review2012/com\\_2012\\_11\\_en.pdf](http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf).
- [3] LI N, ZHANG N, DAS S K, et al. Privacy preservation in wireless sensor networks: a state-of-the-art survey[J]. *Ad Hoc Networks*, 2009, 7(8): 1501-1514.
- [4] DING X, ZHANG L, WAN Z, et al. A brief survey on de-anonymization attacks in online social networks[C]// 2010 International Conference on Computational Aspects of Social Networks (CAsoN), IEEE. c2010: 611-615.
- [5] BACKSTROM L, DWORK C, KLEINBERG J. Wherefore art thou r3579x? anonymized social networks, hidden patterns, and structural steganography[C]//International Conference on World Wide Web. c2007: 181-190.
- [6] NARAYANAN A, SHMATIKOV V. De-anonymizing social networks[C]// 2009 30th IEEE Symposium on Security and Privacy, IEEE. c2009: 173-187.
- [7] WONDRAČEK G, HOLZ T, KIRDA E, et al. A practical attack to de-anonymize social network users[C]//2010 IEEE Symposium on Security and Privacy (SP). c2010: 223-238.
- [8] HAY M, MIKLAU G, JENSEN D, et al. Resisting structural re-identification in anonymized social networks[J]. *The VLDB Endowment*, 2008, 1(1): 102-114.
- [9] NARAYANAN A, SHMATIKOV V. Robust de-anonymization of large sparse datasets[C]// 2008 IEEE Symposium on Security and Privacy(SP). c2008: 111-125.
- [10] NAVEED M, AYDAY E, CLAYTON E W, et al. Privacy in the genomic era[J]. *ACM Computing Surveys (CSUR)*, 2015, 48(1): 6.
- [11] TRAMÈR F, HUANG Z, HUBAUX J P, et al. Differential privacy with bounded priors: reconciling utility and privacy in genome-wide association studies[C]// The 22nd ACM Sigsac Conference on Com-

- puter and Communications Security. c2015: 1286-1297.
- [12] JOHNSON A, SHMATIKOV V. Privacy-preserving data exploration in genome-wide association studies[C]//The 19th ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. c2013: 1079-1087.
- [13] LI N, QARDAJI W, SU D, et al. Membership privacy: a unifying framework for privacy definitions[C]//The 2013 ACM Sigsac Conference on Computer & communications security. c2013: 889-900.
- [14] 刘华玲, 郑建国, 孙辞海. 基于贪心扰动的社交网络隐私保护研究[J]. 电子学报, 2013(8):1586-1591.
- LIU H L, ZHENG J G, SUN C H. Privacy preserving in social networks based on greedy perturbation[J]. Journal of Electronics, 2013, 08: 1586-1591.
- [15] 刘华玲, 郑建国, 孙辞海. 社交网络隐私保护中的随机算法[J]. 信息与控制, 2012(4):197-201.
- LIU HL, ZHENG J G, SUN C H. Randomized algorithm for privacy preservation in social networks[J]. Information and Control, 2012(4): 197-201.
- [16] 吕志泉, 洪澄, 张敏, 等. 面向社交网络的隐私保护方案[J]. 通信学报, 2014(8):23-32.
- LV Z Q, HONG C, ZHANG M, et al. Privacy-preserving scheme for social networks[J]. Journal of Communications, 2014(8):23-32.
- [17] KIDO H, YANAGISAWA Y, SATOH T. An anonymous communication technique using dummies for location-based services[C]//2005 International Conference on Pervasive Services(ICPS'05). c2005: 88-97.
- [18] MOKBEL M F. Privacy in location-based services: state-of-the-art and research directions[C]//2007 International Conference on Mobile Data Management. c2007: 228-228.
- [19] DOMINGO-FERRER J, SRAMKA M, TRUJILLO-RASÚA R. Privacy-preserving publication of trajectories using microaggregation[C]//The 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS. c2010: 26-33.
- [20] GRUTESER M, LIU X. Protecting privacy in continuous location-tracking applications[J]. IEEE Security & Privacy, 2004 (2): 28-34.
- [21] DWORK C. Differential privacy[M]//Automata, languages and programming. Berlin Heidelberg: Springer, 2006: 1-12.
- [22] LEE B, OH J, YU H, et al. Protecting location privacy using location semantics[C]//The 17th ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. c2011: 1289-1297.
- [23] ANDRÉS M E, BORDENABE N E, CHATZIKOKOLAKIS K, et al. Geo-indistinguishability: differential privacy for location-based systems[C]//2013 ACM Sigsac Conference on Computer & Communications Security. c2013: 901-914.
- [24] CHATZIKOKOLAKIS K, ANDRÉS M E, BORDENABE N E, et al. Broadening the scope of differential privacy using metrics[C]//Privacy Enhancing Technologies. c2013: 82-102.
- [25] CHATZIKOKOLAKIS K, PALAMIDESSI C, STRONATI M. A predictive differentially-private mechanism for mobility traces[C]//Privacy Enhancing Technologies. c2014: 21-41.
- [26] XIAO Y, XIONG L. Protecting locations with differential privacy under temporal correlations[C]//The 22nd ACM Sigsac Conference on Computer and Communications Security. c2015: 1298-1309.
- [27] BORDENABE N E, CHATZIKOKOLAKIS K, PALAMIDESSI C. Optimal geo-indistinguishable mechanisms for location privacy[C]//2014 ACM Sigsac Conference on Computer and Communications Security. c2014: 251-262.
- [28] 霍峥, 孟小峰, 黄毅. PrivateCheckIn: 一种移动社交网络中的轨迹隐私保护方法[J]. 计算机学报, 2013(4): 716-726.
- HUO Z, MENG X F, HUANG Y. PrivateCheckIn: trajectory privacy-preserving for check-in services in MSNS[J]. Chinese Journal of Computers, 2013(4):716-726.
- [29] 赵婧, 张渊, 李兴华, 等. 基于轨迹频率抑制的轨迹隐私保护方法[J]. 2014(10): 2096-2106.
- ZHAO J, ZHANG Y, LI X H, et al. A trajectory privacy protection approach via trajectory frequency suppression[J]. Chinese Journal of Computers, 2014(10):2096-2106.
- [30] 吴英杰, 唐庆明, 倪巍伟, 等. 基于聚类杂交的隐私保护轨迹数发布算法[J]. 计算机研究与发展, 2011(5): 578-593.
- WU Y J, TANG Q M, NI W W, et al. A clustering hybrid based algorithm for privacy preserving trajectory data publishing[J]. Journal of Computer Research and Development, 2011(5):578-593.
- [31] AGRAWAL R, SRIKANT R. Privacy-preserving data mining[J]. ACM Sigmod Record, 2000, 29(2): 439-450.
- [32] OLIVEIRA S R M, ZAIANE O R. Privacy preserving frequent itemset mining[C]//The IEEE international Conference on Privacy, Security and Data Mining. c2002: 43-54.
- [33] SAYGIN Y, VERYKIOS V S, ELMAGARMID A K. Privacy preserving association rule mining[C]//International Workshop on Research Issues in Data Engineering: Engineering E-commerce/E-business Systems (RIDE-2EC). c2002: 151-158.
- [34] SAMARATI P, SWEENEY L. Generalizing data to provide anonymity when disclosing information[C]//The 17th ACM Sigact-sigmod-sigart Symposium on Principles of Database Systems. c1998:188.
- [35] SWEENEY L.  $K$ -anonymity: A model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(05): 557-570.
- [36] MACHANAVAJJHALA A, KIFER D, GEHRKE J, et al.  $L$ -diversity: privacy beyond  $k$ -anonymity[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2007, 1(1): 3.
- [37] LI N, LI T, VENKATASUBRAMANIAN S.  $T$ -closeness: privacy beyond  $k$ -anonymity and  $L$ -diversity[C]//IEEE 23rd International Conference on Data Engineering(ICDE). c2007: 106-115.
- [38] ZHANG Y, MAO Y, ZHANG H, et al. Privacy preserving market schemes for mobile sensing[C]//2015 44th International Conference

- on Parallel Processing (ICPP).c2015: 909-918.
- [39] HUA J, TANG A, ZHONG S. Advertiser and publisher-centric privacy aware online behavioral advertising[C]//IEEE 35th International Conference on Distributed Computing Systems (ICDCS). c2015: 298-307.
- [40] HUA J, TANG A, FANG Y, et al. Privacy-preserving utility verification of the data published by non-interactive differentially private mechanisms[J]. IEEE Transactions on Information Forensics & Security, 2016.
- [41] CANNY J. Collaborative filtering with privacy[C]//2002 IEEE Symposium on Security and Privacy. c2002: 45-57.
- [42] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009 (8): 30-37.
- [43] HUA J, XIA C, ZHONG S. Differentially private matrix factorization[C]//The 24th International Conference on Artificial Intelligence(IJCAI). c2015: 1763-1770.
- [44] CAO X, CONG G, JENSEN C S. Mining significant semantic locations from GPS data[J]. The VLDB Endowment, 2010, 3(1/2): 1009-1020.
- [45] ZHENG Y, ZHANG L, XIE X, et al. Mining interesting locations and travel sequences from GPS trajectories[C]//The 18th International Conference on World Wide Web. c2009: 791-800.
- [46] CLARKE R. Person location and person tracking-technologies, risks and policy implications[J]. Information Technology & People, 2001, 14(2): 206-231.
- [47] PELLETIER M P, TREPANIER M, MORENCY C. Smart card data use in public transit: a literature review[J]. Transportation Research Part C: Emerging Technologies, 2011, 19(4): 557-568.
- [48] HUA J, GAO Y, ZHONG S. Differentially private publication of general time-serial trajectory data[C]//2015 IEEE Conference on Computer Communications (INFOCOM). c2015: 549-557.
- [49] MAO Y, CHEN T, ZHANG Y, et al. Protecting location information in collaborative sensing of cognitive radio networks[C]//The 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems. c2015: 219-226.
- [50] ZHONG S, YAO H. Towards cheat-proof cooperative relay for cognitive radio networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(9): 2442-2451.
- [51] CHEN T, ZHONG S. Truthful auctions for continuous spectrum with variable bandwidths[J]. IEEE Transactions on Wireless Communications, 2014, 13(2): 1116-1128.
- [52] HUA J, DU S, ZHONG S. Towards attack-resistant peer-assisted indoor localization[M]//Computer Security—ESORICS 2015. Springer International Publishing, 2015: 417-437.
- [53] HERRING R, HOFLEITNER A, WORK D, et al. Mobile millenium-participatory traffic estimation using mobile phones[C]//CPS Forum, Cyber-Physical Systems Week. c2009.
- [54] THIAGARAJAN A, RAVINDRANATH L, LACURTS K, et al. VTrack: accurate, energy-aware road traffic delay estimation using mobile phones[C]//The 7th ACM Conference on Embedded Networked Sensor Systems. c2009: 85-98.
- [55] DAS T, MOHAN P, PADMANABHAN V N, et al. PRISM: platform for remote sensing using smartphones[C]//The 8th International Conference on Mobile Systems, Applications, and Services, ACM. c2010: 63-76.
- [56] RANA R K, CHOU C T, KANHERE S S, et al. Ear-phone: an end-to-end participatory urban noise mapping system[C]//The 9th ACM International Conference on Information Processing in Sensor Networks.c2010: 105-116.
- [57] BAO X, CHOUDHURY R R. Movi: mobile phone based video highlights via collaborative sensing[C]//The 8th International Conference on Mobile Systems, Applications, and Services, ACM. c2010: 357-370.
- [58] ZHANG Y, CHEN Q, ZHONG S. Efficient and privacy-preserving min and k-th min computations in mobile sensing systems[J].IEEE Transactions on Dependable and Secure Computing, 2015.
- [59] LI Q, CAO G. Efficient and privacy-preserving data aggregation in mobile sensing[C]//2012 20th IEEE International Conference on Network Protocols (ICNP). c2012: 1-10.
- [60] ZHANG Y, CHEN Q, ZHONG S. Privacy-preserving data aggregation in mobile phone sensing[J].IEEE Transactions on Information Forensics and Security, 2016.
- [61] ZHANG Y, ZHANG H, TANG S, et al. Designing secure and dependable mobile sensing mechanisms with revenue guarantees[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(1): 100-113.
- [62] HAO Z, MAO Y, ZHONG S, et al. Toward wireless security without computational assumptions—oblivious transfer based on wireless channel characteristics[J]. IEEE Transactions on Computers, 2014, 63(6): 1580-1593.
- [63] DU S, HUA J, GAO Y, et al. EV-linker: mapping eavesdropped Wi-Fi packets to individuals via electronic and visual signal matching[J]. Journal of Computer and System Sciences, 2016, 82(1): 156-172.
- [64] ZHONG S, ZHANG Y. How to select optimal gateway in multi-domain wireless networks: alternative solutions without learning[J]. IEEE Transactions on Wireless Communications, 2013, 12(11): 5620-5630.
- [65] KATTA N P, REXFORD J, WALKER D. Incremental consistent updates[C]//ACM Sigcomm Workshop on Hot Topics in Software Defined Networking. c2013:49-54.
- [66] MAHAJAN R, WATTENHOFER R. On consistent updates in software defined networks[C]//ACM Sigcomm Workshop on Hot Topics in Software Defined Networking. c2013: 20.
- [67] MCGEER R. A safe, efficient update protocol for OpenFlow networks[C]//ACM Sigcomm Workshop on Hot Topics in Software

Defined Networking. c2012: 61-66.

- [68] REITBLATT M, FOSTER N, REXFORD J, et al. Abstractions for network update[J].ACM Sigcomm Computer Communication Review, 2012, 42(4):323-334.
- [69] HUA J, GE X, ZHONG S. Foun: a flow-ordered consistent update mechanism for software-defined networking in adversarial settings[C]//IEEE International Conference on Computer Communications (INFOCOM).c2016.
- [70] JAIN S, KUMAR A, MANDAL S, et al. B4: experience with a globally-deployed software defined WAN[J]. ACM Sigcomm Computer Communication Review, 2013, 43(4): 3-14.
- [71] CHEN Q, QIAN C, ZHONG S. Privacy-preserving cross-domain routing optimization--a cryptographic approach[C]// IEEE International Conference on Network Protocols(ICNP). c2015:356-365.

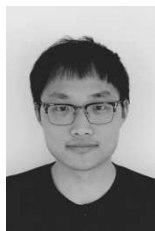
#### 作者简介:



全伟 (1991-), 男, 江苏睢宁人, 南京大学硕士生, 主要研究方向为隐私保护与机制设计。



陈庆军 (1991-), 男, 重庆人, 南京大学硕士生, 主要研究方向为软件定义网络、移动感知中的隐私保护。



王彬入 (1993-), 男, 江苏盐城人, 南京大学硕士生, 主要研究方向为隐私保护和安全博弈。



张保佳 (1991-), 男, 江苏沭阳人, 南京大学硕士生, 主要研究方向为数据完整性验证。



毛云龙 (1990-), 男, 吉林安图人, 南京大学博士生, 主要研究方向为计算机网络、安全与隐私保护、无线网络。



仲盛 (1974-), 男, 江苏南京人, 博士, 南京大学教授、博士生导师, 主要研究方向为密码学、博弈论及其在计算机网络、分布式系统中的应用。