文献阅读报告

一、文献信息

作者：Huaxin Li，Haojin Zhu，Suguo Du，Xiaohua Liang，and Xuemin（Sherman）Shen.

论文题目：Privacy Leakage of Location Sharing in Mobile Social Networks：Attacks and Defense（移动社交网络中位置共享的隐私泄露:攻击与防御）

发表途径：IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING

发表时间：JULY/AUGUST 2018

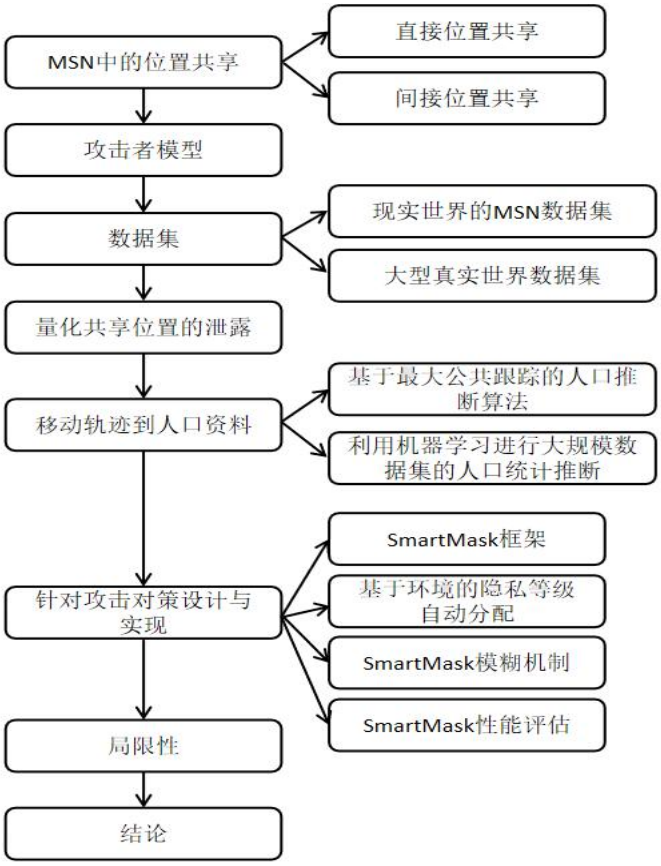
二、问题意义

研究问题：移动社交网络中位置共享的隐私泄露:攻击与防御。建立攻击模型，提出了应对攻击的策略 SmartMask，这是一种基于环境的系统级隐私保护解决方案，旨在自动了解用户在不同环境下的隐私偏好，并为 MSN 用户提供透明的隐私控制。

研究背景：移动社交网络(MSNs)越来越受欢迎，用户可以通过移动社交应用持续感知自己的位置和社会环境，并获得准确和高质量的基于位置的个性化服务。现有的研究工作表明，直接的位置共享机制对于吸引注意力、促进自我呈现、促进和维持社会资本是有效的。但是随着移动社交网络(MSNs)的流行，由于用户位置暴露而导致的隐私泄露的危险也越来越大。

研究意义：通过匹配用户共享的位置和他们真实的移动轨迹，向量化 MSNs 的位置隐私泄露迈出了第一步。为用户提供一种细粒度的、用户透明的隐私控制机制，期望以用户友好的方式为不同的环境提供不同的隐私控制。

三、思路方法



本文中考虑了两种 MSNs 中的位置共享：在 MSNs 中直接共享位置和 MSNs 中的间接位置共享。

建立了攻击者模型。攻击者的目标是建立目标受害者的位置概况，并推断出目标的敏感数据或人口统计数据。文章考虑了两种对应于两种位置共享方法的攻击者：随意跟踪攻击者和连续跟踪攻击者。

选取了两个数据集。第一个数据集是一个真实的 MSN 数据集，数据集为我们招募的 30 名志愿者，来自不同部门和不同等级校园，收集他们直接从微博和人人网分享位置(直接共享跟踪),间接从微信分享位置,momo 和 Skout(间接共享跟踪),流动痕迹以固定时间间隔(地面实况跟踪)。第二个数据集是一组大规模的真实世界 Wi-Fi 流量记录，涉及 5 个月内 22843 个用户的数据。该数据集包含校园内 98 个 Wi-Fi 热点的 MSNs 流量日志，MSNs 流量日志记录位置轨迹。每个日志包含匿名的用户 id、MSN 名、位置和访问时间。同时，该数据集还提供了匿名的用户属性，如性别、教育程度等。

为了获取移动用户移动的可预测性程度，考虑一个基于熵的定义来建模用户移动模式。为了很好地量化共享位置的泄露,提出了两个新的度量来量化共享位置和实际移动模式之间的相似性，这表明了从位置共享中泄露了多少隐私。

研究了对手是否可以利用这些不完整的信息来推断用户的人口统计信息。介绍根据不同

规模的数据集推断人口统计属性的不同方法。基于最大公共跟踪的人口推断算法，基于共追踪的方法开始，在小范围内推断目标用户的人口统计特征。在此基础上，提出了一种更具可扩展性的机器学习方法——利用机器学习进行大规模数据集的人口统计推断，并将其应用于大规模数据集的数据集 II 上进行人口统计推断。最后，比较了这两种方法的优点，并讨论了它们的应用。第一种方法的性能优于机器学习方法。但是在相同的数据量和运行环境下，基于共道的方法的运行时间几乎是基于机器学习的方法的 496 倍。因此，我们在不同的情况下提出了不同的方法。

提出应对攻击提出的对策——开发 SmartMask，这是一种新颖的隐私保护框架，旨在为 MSN 用户提供细粒度的隐私管理。SmartMask 的主要思想是通过根据不同的位置和用户的偏好设置隐私等级，来平衡隐私保护和实用功能之间的权衡。其具有的特性：环境驱动的隐私管理、细粒度的位置隐私控制、隐私级别自动分配。

SmartMask 由四个组件组成：环境生成器、保密级别生成器、指定的接口、模糊处理引擎。

SmartMask 应当具有基于环境的隐私等级自动分配功能，因此文章提出了一种基于决策树的自动隐私级分配方法。

对于 SmartMask 的模糊机制，文章实现了一种基于两种不同的混淆技术的混合混淆方法。当隐私等级较低或中等时，SmartMask 通过随机组合这些模糊操作符来进行模糊处理(低隐私等级比中等模糊程度轻)。在隐私等级高的情况下，用户不想在这些隐私等级中暴露他们的真实位置。因此，SmartMask 采用了一种简单的伪装策略，使模糊结果偏离到最近的公共区域和较不敏感的区域。此技术对于防御恶意或不需要的位置请求非常有用。

为了评估 SmartMask 的性能和优点，又进行了一次为期 3 周的实验。考虑反映用户 N 个最私密位置的 N 个覆盖率、考虑相对熵度量、进行敏感性分析。评估效用的减少，评估人口保存的效果，测量时间延迟评估性能，测量不同情况下的电池消耗率评估能耗。

为了无法访问移动社交网络中的大规模数据集,招募志愿者来执行实际实验收集共享位置,造成数据集有一个更大的数据集虽然这种限制可能会导致数据集存在偏差，它不会使通过共享位置进行的隐私推断无效。作为未来的工作之一，将考虑一个更机智的对手，可以在移动社交网络中收集大规模的共享位置,以更好地了解不同用户对位置共享导致的隐私泄露的影响。

四、实验结论

位置共享在 MSNs 中的普及引起了越来越多的隐私问题。在这项工作中，我们根据实际收集的数据集对共享位置和真实位置的相似性进行了定量评估。我们的定量评估显示,尽管直接位置共享和间接位置共享只有揭示 16 和 33%的用户的真实利益点(POIs),攻击者可以利用痕迹在不同用户之间的相似性来推断其年龄、职业、生活的地方,性别和教育程度的成功

率,69.2,53.8,54.5,76%和 73。然后，我们提出了 SmartMask，这是一种系统级的解决方案，可以在不显著降低服务质量的情况下阻止位置隐私泄露。SmartMask 可以根据位置环境自动学习和生成位置的隐私级别。作为一个通用平台，SmartMask 可以结合其他先进的模糊处理技术来解决更广泛的位置隐私问题。

五、启发思考

移动社交网络(MSNs)如 Facebook、微博、微信、陌陌等越来越受欢迎。其中许多 MSN 支持直接位置共享如位置签到、地理定位标签，还有一些社交软件提供了一种隐式的位置共享方式，只显示粗略的邻近信息，如“小明在三公里以内”、“周围的人”。这些直接或者间接的位置共享增加了许多的用户位置安全隐患。

该论文便是基于以上问题背景而研究的基于移动社交网络中位置共享的隐私泄露的攻击与防御。针对社交软件的位置隐私问题，已有不少研究。然而，由于 MSN 上的位置共享而导致的隐私信息泄露问题的量化却很少被关注。

该文章创造性的量化共享位置和实际移动模式之间的相似性，来反映从位置共享中泄露了多少隐私。提出了 SmartMask——一种新颖的隐私保护框架，旨在为 MSN 用户提供细粒度的隐私管理。并且通过实验验证 SmartMask 的性能。

论文由浅入深，层层引入，初读便觉得有条有理，并不晦涩难懂。内容与实际生活联系紧密，尤其话题是现在人们离不开的社交软件以及用户隐私问题，这更加增加了阅读兴趣。通过阅读这篇文献，我不仅得到了知识，更加了解了移动社交网络中位置共享的隐私泄露，也学到了研究问题的思路，以及许多研究过程中的方法，获得了很大的启发。