ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.21, Supplement, December 2010, pp.298−309

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

移动社交网络中的隐私设计*

谈 嵘+, 顾君忠, 杨 静, 林 欣, 陈 鹏, 乔哲峰

(华东师范大学 计算机科学与技术系,上海 200241)

Designs of Privacy Protection in Location-Aware Mobile Social Networking Applications

TAN Rong+, GU Jun-Zhong, YANG Jing, LIN Xin, CHEN Peng, QIAO Zhe-Feng

(Department of Computer Science and Technology, East China Normal University, Shanghai 200241, China)

+ Corresponding author: rtan@ica.stc.sh.cn, http://www.ecnu.edu.cn

Tan R, Gu JZ, Yang J, Lin X, Chen P, Qiao ZF. Designs of privacy protection in location-aware mobile social networking applications. ***Journal of Software***, **2010,21(Suppl.):298−309.** http://www.jos.org.cn/1000-9825/10031.htm

Abstract: The emergence of location-aware mobile social networking applications (SNAs) has gained considerable prominence in ubiquitous computing. However, the weak designs of privacy protection result in more potential risks that users have to face. This paper presents a deep insight into the designs of privacy protection, especially from the perspective of location privacy. Three main potential risks and some necessary privacy-related functionality in current applications are analyzed and summarized. And with respect to the risks and weak protection methods, we propose the designs of privacy protection in LaMOC which is a location-aware mobile cooperation system developed by us. The designs include both the policy-based mechanisms and computational approaches to protect users' location privacy. Policy-based mechanisms enable users to fully control their locations disclosure while computational approaches can protect users' real locations from malicious attackers. The designs provide a reasonable level of privacy protection in location-aware mobile SNAs to some extent.

Key words: privacy protection; mobile social networking; location awareness; mobile cooperation; social relationship

摘 要: 对现有移动社交网络应用中的位置隐私问题提出了新的理解,总结了 3 类主要可能造成隐私泄露风险的原因和 3 类与之对应的必要保护手段,并且提出了一种结合策略与算法的位置隐私保护机制.在同时构建的基于位置感知的移动协作应用原型系统 LaMOC 上,该隐私保护设计能够有效从两方面全面保护用户的位置隐私.一是支持用户通过灵活的策略方式,完全控制自身的位置信息共享方式;二是通过隐私保护算法使恶意攻击者无法得到用户的敏感位置信息.该隐私保护设计在一定程度上有效保护了移动社交网络应用中的用户位置隐私.

关键词: 隐私保护;移动社交网络;位置感知;移动协作;社会关系

过去数年来,基于位置感知的移动社交网络应用(location-aware mobile social networking applications)在普

* Supported by the National Natural Science Foundation of China under Grant No.60903169 (国家自然科学基金); the Project of Science and Technology Commission of Shanghai Municipality of China under Grand Nos.09510703000, 10dz1500103 (上海市科委科研计划项目)

Received 2010-07-01; Accepted 2010-12-10

适计算研究领域取得了长足的发展.与传统桌面社交网络不同,移动社交网络应用将用户的位置信息同实际应用结合在了一起.现有的应用如 Dodgeball[1]和 Loopt[2]都在用户界面中提供了一张用户当前位置的电子地图,在这之上用户不仅可以知道自己所处的位置,而且可以随时发现周围的好友,同时还能将自己此时此刻在生活中所拍的照片与所写的博客发布在地图上的某个确切地点与他人分享.这使得移动社交网络应用比传统桌面社交网络更加能够促进同一个城市中人与人之间的交流,受到了相当一部分年轻人的喜爱.Loopt 早已宣布其用户数量已经突破 10 万,并且还在快速增长中.在中国,移动社交网络应用同样也发展迅速,贝多[3]截止到 2008 年底也已经拥有了超过 2 万的用户.

然而经历了快速的发展阶段之后,基于位置感知的移动社交网络应用也面临了一系列巨大的挑战.安装在用户手机上的该类应用,无时无刻不在将用户的当前位置信息泄露出去,这使得如何保护用户的位置隐私成为了这其中的关键问题之一.一方面,用户将自己的位置信息共享,可以增加同好友之间的联系,也可以结交更多的朋友.另一方面,用户则必须承担被恶意跟踪者跟踪的风险.如何采取适当措施保护用户的位置隐私,平衡共享位置信息的利与弊是每个移动社交网络应用需要解决的问题.

在美国,政府为保护儿童的信息隐私制定了相关法律 COPPA[4](Children's Online Privacy Protection Act).欧洲也同样实行了与其类似的法律[5].各国的研究学者们也在各方面进行了探索与研究,努力提高移动社交网络应用的隐私保护能力.最近几年,对移动社交网络应用中的位置隐私问题研究主要分为以下两类:

1) 提出新的位置隐私保护机制,并建立应用原型.文献[6]建立了一个利用 WIFI 网络定位的应用原型,用户可以为个人信息设置不同的隐私等级.文献[7]提出了一种可以加密的通信协议,以此保证通信过程中用户互相之间的隐私安全.文献[8]将语义网与移动网络相结合,提出了一种利用信任机制保护用户隐私的方式.

2) 对共享位置信息问题的本身进行探索,为其他研究学者提出保护策略提供借鉴.文献[9]制定了一套实验问卷,并利用经济学与心理学方法,探究用户的位置信息的商品价值.文献[10–13]针对不同场景下,用户共享位置信息的意愿进行了调查.统计结果显示,不同对象、时间与环境等因素下,用户的意愿都会有区别,这也从另一个侧面反映出该问题的复杂性.文献[14]利用 Bellotti 和 Sellen 的反馈与控制模型对现存的移动社交网络应用进行了分析比较,总结了它们各自存在的一些不足.

可以看出,位置隐私保护问题已经引起了国内外研究学者们的高度重视.本文对移动社交网络中的位置隐私问题进行了分析与总结,提出了自己的理解,并设计了一套相应的位置隐私保护机制.首先,通过对 4 个位置隐私关键问题的回答,深入分析了移动社交网络中位置隐私的根本问题.其次,对现有主要移动社交网络应用中的隐私保护手段进行分析,总结了 3 类可能造成隐私泄露风险的原因,并提出 3 种必要的位置隐私保护策略.最后,提出了一种结合策略与算法的隐私保护机制,并在应用原型系统 LaMOC[15](location-aware mobile cooperative system)上实现.策略机制通过提供用户更加灵活的控制手段,使用户可以自己调整位置隐私共享策略,弥补了绝大多数移动社交网络应用中用户可使用的保护手段相对简单这一不足.算法机制从用户位置信息的存储与访问角度出发,首先对位置信息采用不同粒度进行存储.数据库中不再存储用户的精确位置信息,取而代之是经过模糊处理后的圆形区域或矩形区域,有效防止了恶意用户获取用户的精确位置信息.其次,针对应用场景如何处理模糊后的位置信息问题,以好友推荐[16]与好友集结两个应用场景为例,提出了基于 Skyline 的改进算法,为今后新的应用场景提供了借鉴.

本文第 1 节对移动社交网络中的位置隐私问题进行分析.第 2 节总结 3 类潜在风险以及相对应的必要保护手段.第 3 节提出结合策略与算法的位置隐私保护机制.最后总结全文,并展望未来的研究工作.

1 理解位置隐私问题

隐私保护的设计问题向来都是普适计算研究领域中的一个重要研究课题[17,18].对于移动社交网络来说,除了对用户基本个人信息进行保护外,对用户的位置信息保护也同样重要.在不影响用户使用体验的基础上,如何为其提供一套合理的隐私保护机制,一直是移动社交网络开发者需要切实考虑的问题.只有当开发者对位置隐私的问题有了深刻的理解,才能为用户提供最为完善的隐私保护设计.这里提出了 4 个与位置隐私相关的核心

问题,通过回答这 4 个问题,可以帮助设计人员从不同角度理解移动社交网络中的位置隐私问题,从而更好地完善隐私设计.这 4 个问题分别是:

1) 为什么需要分享位置信息?

在平时生活中,“你在哪里?”、“你要去哪儿?”总是出现频率最高的问题之一.父母以此确保孩子的安全,老板以此检查员工的工作情况,朋友之间以此嘘寒问暖,增进感情.可以说,潜意识里人与人之间对彼此的位置信息非常关心.实际上,位置信息经常能够反映出人们当前的实时状态,例如是否在家休息、办公室上班或是医院看病等等.

移动社交网络中,分享彼此的位置信息也能够促进本地间人与人的交流.21 世纪人类生活方式已经发生了翻天覆地的变化,随着城市规模的逐渐扩大,人与人之间的情感联系反而变得越来越生疏.究其原因之一在于人们缺少与他人沟通交流的机会,人们很难方便地找到与自己志同道合的朋友.进化心理学家指出,人类只有通过彼此紧密的联系才能发展.所以说,在移动社交网络中分享自己的位置,以此来寻找本地的好朋友,不失为一个增加人与人沟通交流机会的好方法.

另一方面,分享位置信息也丰富了各种计算机服务,使更多应用场景的实现成为可能.比如通过基于位置的服务(location-based services)中的定位功能,人们不用再担心在陌生的城市迷路.同样,人们不用再为了决定一个聚餐吃饭的地方,反复与不同好友打电话,只需要根据每个人的具体位置和他们的偏好,就可以让计算机来帮助你选择一系列合适的餐馆.分享位置信息使计算机支持的协同工作(computer supported cooperative work,简称 CSCW)更为简单.

2) 分享位置信息的风险是什么?

虽然分享位置信息的好处是显而易见的,但是有时候它也可能为分享者带来麻烦,甚至危险.第一,用户无法再对自己的实际位置进行掩饰.一旦移动社交网络应用所暴露的用户当前位置与其所言不符,则会使用户陷于尴尬的境地.第二,用户可能被他人跟踪.据国外媒体报道,曾有人利用装有 GPS 的手机跟踪自己的女友,监视其一举一动,这种行为极大地侵犯了他人的个人隐私.而这种风险也确实可能发生在移动社交网络中,因为一旦分享自己的位置信息,随时可能有人通过你所公开的位置信息来监视你.第三,位置信息其本身虽然只是数字,但是它的泄露也可能导致其他个人隐私的泄露.比如,恶意的商业竞争对手可以通过你分享的位置信息,掌握你与客户见面会谈的时间地点,造成商业机密外泄;窥探者也可以通过调查跟踪目标人物每一天的行程,对其日常生活习惯等个人私生活进行了解,引起不必要的个人隐私泄露.

3) 风险是怎么发生的?

大多数情况下,风险的发生只有两种原因:第一,移动社交网络没有提供用户足够的位置信息保护手段,使用户无法切实保护自己的位置信息.虽然文献[12]显示,影响用户分享位置信息意愿的主要因素是询问者与分享者之间的关系,并且绝大多数移动社交网络应用也允许用户自定义哪些其他用户可以得到自己的位置信息,但是文献[10]指出,用户实际的分享意愿也会随不同的地点、活动和心情而经常改变.例如,86%用户愿意对来自同一城市的其他人共享自己的位置信息,而对于远离自己居住地的人,只有 55%的用户愿意将位置信息与其分享.对于用户当前所进行的活动类型对共享意愿的影响,96%的用户愿意在自己做家务的时候分享位置信息.另外,84%用户与 81%用户分别愿意在锻炼与打电话时分享当前位置,而当用户在学习工作或与好友在一起时,该数据则分别只有 63%与 65%.另一个重要因素是心情,当用户比较寂寞时,82%的情况下他们愿意分享自己当前的位置,而当他们处于愤怒状态时,则只有 57%的可能愿意分享.其他心情诸如开心、冷静、有压力以及悲伤,则分别有 77%,77%,72%与 64%的可能性分享自己的位置信息.因此,简单划一的保护手段是无法起到真正的保护作用.

第二,用户自身的疏忽.即便移动社交网络应用能够提供足够的保护手段,也可能因为保护手段的过于繁琐而致使用户的疏忽,进而导致最终的风险产生.文献[13]为用户提供了一套可以自定义的隐私保护策略,用户可以根据不同对象,时间与地点制定相应的位置信息保护规则.研究显示,用户在刚开始时会频繁调整自己的隐私策略,增加隐私保护规则,制定的规则数量也往往超过 8 条,以保证自己的位置信息不被泄露.但是随着时间的推

移,用户会最终停留在某个策略规则集合上,不再进行调整.而且通过分析,用户对自己最开始使用的策略满意度仅有59%,即使是最后一直使用的那个策略,也仅有70%.用户并没有因为自己增加了许多保护规则,而对隐私保护策略感到真正满意.由此可见,复杂的隐私保护策略虽然能够保证用户在各种情况下保护自己的位置信息,但是由于其操作的繁琐与麻烦,最终会使用户选择偷懒的方式,停留在不安全的机制之上.

4) 为什么很少有用户会意识到可能有风险发生?

虽然很多专家在移动社交网络形成之初就曾指出,用户分享自身的位置信息可能会面临不可预见的风险.但是实际情况却显示,用户并不认为分享位置信息会为他们带来问题.文献[9]利用经济学与心理学的研究方式,试图评估用户对自己位置信息的心理价值的衡量标准.结果显示,有 11%的被调查者并不在意自己的位置信息被他人获取,仅需 1 英镑的价格,他们就允许某段时间内其具体位置信息被他人监视.而大约一半的人则更愿意以 10 英镑的价格将自己某段时间的位置信息向其他人分享.只有极少数人认为自己位置信的价值应该超过400 英镑.文献[19]也指出,人们并不需要或者期望能在与外界接触的过程中得到全面的隐私保护.然而,人们对隐私保护意识的缺乏并不意味着这样的风险并不存在.

Ropeik[20]等人在他们的《风险:是发现生活中安全与危险的实用指南》一书中列举了一些关于人们习惯于忽视自己身边所存在风险的理由,其中有两条非常合适用来解释移动社交网络中用户风险意识淡薄的原因.第一,绝大多数人会轻视那些可能对自己产生利益的风险.第二,绝大多数人会轻视那些他们认为能够驾驭的风险.事实上,在本节第一个问题中已经列举了分享位置信息可能为用户带来的好处.用户在与他人分享自己的位置信息时,从自身角度出发,更多的是出于积极的目的,也因此他们会在潜意识里轻视了由此可能导致的风险.另一方面,移动社交网络应用所提供的单薄的隐私保护手段,也使他们错误地估计了对自己位置信息的控制力,进一步导致风险意识的降低.

以上 4 个不同角度的分析探讨,帮助我们深入理解了位置隐私问题的实质,同时也进一步突显出研究位置隐私保护的必要性.

2 引起风险的主要原因与对应的关键保护机制

第 1 节对移动社交网络中的位置隐私进行了详细的分析,预防风险的发生除了需要用户自身提高警觉外,更最重要的是应用设计者能够为用户提供灵活全面的隐私保护机制.但是实际使用中,现有的基于位置感知的移动社交网络应用并没有提供足够的保护手段.本节对一款现有的基于位置感知的移动社交网络应用软件进行分析,在此基础上,总结了该类应用软件中造成隐私信息泄露风险的 3 种主要因素,同时提出了与之对应的关键保护机制.

2.1 研究案例——贝多

贝多是国内最早一批进入移动社交网络领域的软件,其用户数量在 2008 年底就达到了 2 万以上.本文随机选取了 400 位贝多的用户,并对他们进行了调查.图 1 显示,同国外桌面社交网络 Facebook 用户年龄结构分布相对均匀相比[21],贝多的用户群年龄层主要分布于 18 岁~34 岁之间,并主要以年轻人为主,这和贝多将青少年与都市白领定义为其目标群体基本符合.同时,图 1 也显示了国内的桌面社交网站使用者的年龄结构分布更趋于年轻化.由此可见,同国外相比,国内社交网络使用者的年龄层相对较低,他们对新事物的接受能力较强,但是风险防范意识则相对较弱.

贝多允许用户彼此互相共享位置信息,在电子地图上,用户的即时位置使用一个笑脸图标标识,其他愿意分享位置信息的用户则用小人图标表示.贝多同时支持用户将当前的照片或者博客与特定的地点相关联,与同其他用户分享,这些照片或博客在电子地图上使用星状图标进行标记.与传统的即时通信软件一样,每一个贝多用户都能够建立一组好友列表,与他们进行即时消息的交流.另外,用户的一些基本信息如性别、年龄和居住城市也能被其他用户查询得到,以此方便其他用户寻找到合适的朋友.在对 20 位随机用户进行调查后发现,用户好友列表中的绝大多数对象并不是自己现实生活中所熟悉的好友,相反是通过贝多这个虚拟平台认识并添加的.究其原因,一方面,贝多软件并未大量普及,所以他们很多好友并未使用贝多.另一方面,他们希望在一个全新的

虚拟世界中认识其他更多的新朋友,拓展自己的社交圈.

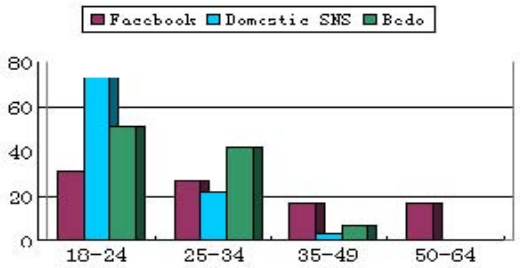


Fig.1 The age distribution of users of Facebook, social networking sites in China and Bedo

图 1 Facebook,国内社交网站与 Bedo 的用户年龄分布

Fig.2 Main screen of Bedo

图 2 贝多主界面

针对贝多中的位置信息共享问题,其使用基站定位方式获得其用户的位置信息,与用户当前精确位置的误差大约在 1 公里左右.用户的即时位置信息会以每 10 分钟一次的频率进行更新.用户除了可以在电子地图上查询得到好友的位置信息外,贝多同时也提供了一套在好友列表上指示好友距离的方法.比如如果该好友当前位置距离用户比较近时,该好友昵称的旁边就会显示一个正在走路小人的图标;而距离相对较远时,该图标则会变化为一辆自行车;若该好友与用户不在同一个城市,则会显示为一架飞机.贝多为其用户提供了 3 种共享位置信息的方式:仅与好友共享,与所有用户共享,不与任何用户共享.

2.2 造成风险的3种原因

2.2.1 保护手段不灵活

虽然贝多为其用户考虑了用户位置信息隐私的保护问题,并为用户提供了 3 种共享位置信息的方式.但是实际使用中,这种不灵活的保护手段无法真正起到保护作用.首先,无法保证用户愿意共享位置信息的对象覆盖了用户的所有好友.即便在好友列表中,用户的共享意愿也会有所区分.一旦用户将共享位置信息的决策设置为“仅与好友分享”,意味着用户必须面临所有好友都能时刻掌握其当前位置信息的风险.其次,无法保证用户位置信息在某些特定情景下不被泄露.这里的特定情景包括,一天中某个特别的时间段、用户处于某个特别的地点以及用户正在进行某个特别的活动.如果隐私保护机制不与特定情景相结合,一旦用户疏忽,位置隐私泄露将不可避免.

2.2.2 缺少反馈机制

反馈机制的最大作用在于对恶意用户起到警示作用,及时提醒正常使用的用户,防范于未然.反馈机制包括两个方面:第一,提醒机制.当其他用户查询用户当前位置信息时,应用应该有义务向被查询的用户发出提醒,询问是否向该对象共享位置信息.如果应用认为太频繁的提醒与询问会造成用户使用感受降低,那么结合共享位置信息的方式,应用应该针对那些在某个时间段中频繁查询其他用户位置信息的用户进行双方面的提醒,以保证用户的位置信息隐私不被恶意用户获取.第二,审计机制.用户应该能够回顾过去某段时间内对其位置信息进行查询的统计信息,包括查询对象、查询时间和查询次数等.用户通过审计机制可以随时了解其他用户针对自己位置信息的查询情况,及时发现不正常的情况,并可以有针对性地调整自己的隐私保护策略.

2.2.3 数据存储方式不合理

贝多中针对用户位置信息的存储机制并不合理,而绝大多数其他同类软件也存在类似问题.首先,由于用户的位置信息会被存储在服务器端,即使用户离线,用户最后时刻的位置信息仍旧能够被其他用户查询得到.事实上,应用不应该对用户的位置信息进行存储,因为一旦服务器遭受攻击,用户的位置隐私遭到泄露,可能造成不可预见的结果.如果应用对用户提供的服务需要对用户的位置信息进行存储,也应该采用一定的保护手段,直接将用户精确历史位置信息进行存储是不安全的.其次,绝大多数移动社交网络会使用第三方的地理信息服务,例如贝多使用的是谷歌地图.服务提供环节的增加,也增加了隐私信息泄露的可能性.最后,许多移动社交网络应用为了方便用户的登录,会将用户的登录号以及密码缓存在本地手机上.另外,包括用户平时使用中的历史聊天记录也会存储在本地.这种做法面临的潜在风险是,如果用户的手机遗失,不仅用户自身的重要信息可能遭到滥用,而且由于其他人可以方便进入系统,对其他用户的信息隐私安全也会造成影响.文献[22]的数据显示,有 22%的手机拥有者曾经遗失过他们的手机.因此,登录号以及登录密码不应该在用户手机本地进行缓存.同时,应用应该和用户的手机号进行绑定,使其他使用者不能够轻易地进入系统,以此保证用户的个人信息隐私安全.

2.3 对应的关键保护机制

针对以上 3 种造成隐私信息泄露风险的因素,移动社交网络应用应该具备与之相对应的关键保护机制.第一,提供用户自定义的位置隐私共享策略.用户应该能够具体决定哪些用户可以查询到其位置信息,而哪些用户不可以.并且用户应该能够设置特定情境下的位置隐私共享方式,包括用户处于哪些特别的时间段和位于哪些特别的地点.用户自定义策略并非意味着用户需要自己编写隐私保护脚本,相反,应用应该提供用户灵活的设置机制,引导帮助用户快速完成隐私策略的设置.第二,应用应该设立反馈机制.在发生不正常的情况下,应用必须能够提醒用户其位置信息可能被其他有恶意的用户过度查询.同时,应用也应该警告恶意用户,阻止其恶意行为的发生.第三,应用应该为用户提供一套合适的拒绝机制.当用户并不希望将自己的位置信息同某些其他用户共享时,直接冰冷的拒绝可能会使用户陷于尴尬的境地.应用应该为用户提供诸如“网络繁忙,请稍后再试.”、“用户暂时不在服务区”和“用户正忙”等说辞帮助用户回答,提高用户的使用体验.上述 3 种关键保护机制是本文认为的所有移动社交网络应用中都应该具备的隐私保护机制,在此之上,如果能够加入更多的信息隐私保护手段,那么就可以更加全面地保护用户的个人隐私,使用户可以放心地使用该类应用软件,促进其健康快速地发展.

3 LaMOC 中的位置隐私保护设计

3.1 LaMOC系统介绍

LaMOC 应用原型系统由华东师范大学计算机应用研究所联合复旦大学媒体计算与 Web 智能研究团队、中国科学院上海技术物理研究所以及国外知名大学共同参与研究与开发,旨在实现一个具备可以随时、随地、为任何人、任何事提供的各种移动信息服务(anyone and anything at anywhere and anytime,简称 4A 服务)的移动计算协作服务平台.LaMOC 系统研究探索了移动计算环境中诸多前沿问题,包括基于位置服务(location-based services)中情景信息的获取与处理、情景信息在有资源限制的移动设备上展示与存储、移动计算环境中的协作模式[23]和隐私保护设计[15]、基于图像的空间位置检索技术[24]、伪卫星 3D 位置服务[25,26]以及用户行为模式

的挖掘与兴趣推荐[27]等等.其中,隐私保护设计针对移动计算环境中基于位置的服务和移动社交网络中的隐私问题,提出了一种隐私保护机制,该机制结合了用户策略与空间信息存储与访问算法,降低了用户信息,特别是用户的位置隐私发生泄漏的风险.

图 3 显示了 LaMOC 系统的软件体系结构,其中用户信息隐私保护机制由 3 个模块负责,分别是登录管理模块、安全管理模块和信任管理模块.登录管理模块负责管理每一个登录系统用户的有效性.安全管理模块不仅为每一个用户提供了一个安全的与系统通信的会话(session)通道,保证了用户与系统之间以及分布在各处的子系统之间的通信安全,而且也保证了情景信息的存储与访问安全,这里的情景信息包含用户的位置信息.信任管理模块保证了 LaMOC 运行在一个可信的计算环境中,为系统中模块与模块之间,用户与信息之间,用户与服务之间建立了完整的信任机制.当有服务对用户的信息进行请求时,它根据用户自定义的信息隐私策略,采用合适的信息存储与访问算法,对隐私信息进行处理后返回该服务请求.其中,本文所研究的位置隐私保护机制正是在信任管理这一模块下的应用.

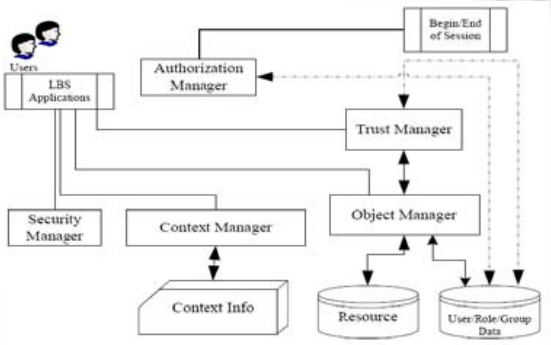


Fig.3 The software architecture of LaMOC

图 3 LaMOC 系统的软件体系结构

3.2 基于用户策略的位置隐私保护机制

隐私保护机制最重要的任务在于为用户提供一套可供其灵活控制的位置信息共享方式.对于普通移动社交网络使用者来说,应用背后的信息处理算法是其不可见的,也是无法理解的.用户更加关心的是应用是否能够提供直观的操作来帮助自己控制信息共享.事实上,隐私保护的目地在某种程度上是为用户在使用应用的过程中提供其安全感.如果用户认为自己在使用应用的过程中,应用提供的信息共享方式能够满足或者保证其个人信息隐私的安全,那么他就会认为这种信息隐私保护策略是合适的,反之则不然.

因此,在设计 LaMOC 系统中基于用户策略的隐私保护机制时,主要提出了两个基本要求:第一,用户可使用的保护策略要足够全面.第二,策略的使用方式必须足够简单与方便.如果策略的使用方式是复杂的,降低了用户的使用体验,不仅不利于保护用户的信息隐私,而且可能阻碍移动社交网络应用的发展.基于这两个要求,LaMOC 系统的保护机制在覆盖了第 2.3 节中所阐述的 3 种关键保护机制以外,还为其用户提供了两种新的保护策略:改进的用户位置显示方式和隐私情景调节模式.

3.2.1 改进的用户位置显示方式

现有的位置感知移动社交网络应用通常使用全球定位系统(GPS)、无线网络(WIFI)或基站定位等方式获取用户的位置信息,不同的获取手段所得到的信息准确度也是不同的.例如,LaMOC 系统使用 GPS 装置,其在实际使用过程中的精确度在几十米至百米之间,用户当前的活动地点可以被很方便地识别出来.相反,使用基站定位方式,其位置信息的精确度则在千米左右.同时,在使用精确定位方式时,用户的实际位置通常使用精确的点来表示;而使用精度相对低的方式时,用户的实际位置则会使用圆形区域来表示,在一定程度上模糊了用户实际位置.来换而言之,随着用户位置信息的精确度越高,其隐私的泄露风险反而越大.因此,灵活改变用户位置的显示方式,也有利于起到保护用户隐私的作用.

一般情况下,LaMOC 使用一个小人图标标记其他用户的当前地理位置.而在实验室内部的测试研究中发

现,用户经常会陷入某种尴尬的境地.例如,用户实际不愿意在此时此刻向好友们共享其精确的位置信息,而当其他用户对其位置进行查询时,系统提供的委婉拒绝方式不足以使其感到满意.用户指出希望得到一种既能够使其他用户可以查询其位置信息,又不会充分暴露其实际位置的方式.虽然使用圆形区域表示用户的位置可以在一定程度上起到模糊用户实际位置的作用,但是对于熟悉该用户情况的好友而言,该方式并不能起到有效保护用户位置隐私的作用.这里,LaMOC 系统提出了一种改进的用户位置显示方式,使用指向性位置表示与语言描述相结合的方式,代替点或圆的位置表现形式.指向性位置表示方式是指,使用一个指向被查询对象当前实际位置方向的箭头来指示该用户的位置.语言描述是指,通过使用 3 种描述语句指示被查询对象与查询者之间的距离.这 3 种语句分别是:“XX 离你很近”、“XX 离你不远”和“XX 离你很远”,其中“XX”表示被查询对象的名称.图 4 中,左下脚箭头即为系统提供的指向性箭头,而在下面的深色框中则显示了语言描述.

Fig.4 The directional arrow and linguistic terms indicates the target user’s location

图 4 结合指向性箭头与语言描述的位置表示方式

3.2.2 隐私情景调节模式

除了隐私策略的全面性外,策略使用方式的简单与方便程度也决定了该策略的保护效果.同桌面应用相比,在移动社交网络应用中应该避免复杂及频繁的操作,要求用户针对不同情况频繁调节其隐私策略显然是不合适的.因此,LaMOC 在参考了即时通信软件中用户状态设置模式的基础上,为其用户提供了 4 种快速方便调节隐私策略的模式,它们分别是:派对(party)模式、亲密(intimacy)模式、普通(normal)模式和不可见(invisible)模式.在派对模式下,用户的位置信息与系统中所有其他用户共享,任何其他用户都可以在其电子地图上看到你当前的位置.利用此模式,用户可以与周围相近的其他用户成为朋友,拓展自己的人际关系圈.亲密模式根据用户预先自定义的隐私访问策略,决定用户好友列表中哪些好友的位置信息访问请求可以得到满足,对于好友列表中被用户屏蔽的其他用户,则使用合适的拒绝机制进行返回.同时,陌生用户无法对用户进行访问请求.普通模式使用改进的用户位置显示方式,当用户的位置信息被其他用户要求访问时,系统会使用指向用户实际位置的箭头与描述距离的语言表示用户位置.此模式下陌生用户同样无法得到访问许可.在不可见模式下,任何其他用户都无法获取用户的位置信息.采用这种隐私情景调节模式,能够避免用户频繁调整其位置信息的共享方式,有效减少了用户的操作次数,同时能够起到保护用户位置隐私的作用,提高了用户的使用体验.

3.3 基于算法的位置隐私保护机制

基于用户策略的隐私保护机制使用户在直观上能够完全控制位置信息的共享方式,增加其使用安全感.而另一方面,则需要通过对后台基于算法的隐私保护机制的改进,从技术上真正实现用户信息隐私的保障.两者相辅相成,缺一不可.在 LaMOC 系统中,基于算法的位置隐私保护机制主要分为位置信息的存储与位置信息的访问两个方面.针对位置信息的存储,系统采用了位置信息的 k-匿名[28](k-anonymity)方法进行了改进.所谓位置信息的 k-匿名是指,如果该用户的位置信息与其他 k–1 个用户的位置信息无法区别开来,那么则称其位置信息满足 k-匿名.所以为了达到 k-匿名的条件,数据库中所存储的用户位置信息,由原来精确的经纬度坐标,被一块矩形区域或圆形区域替代,该区域要求同时包含其他 k–1 个用户的位置信息.如果移动社交网络应用的数据库遭到

恶意攻击,那么攻击者也无法将用户的实际位置信息与其他 k–1 个用户的信息相区分,从而有效保护了用户的位置隐私.

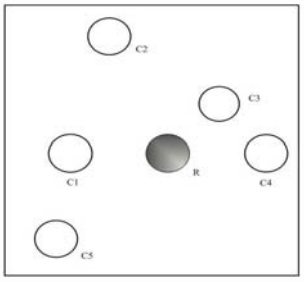
然而,传统的空间数据查询方法都是基于点对点(point-to-point)的,用户的位置信息是确切的点坐标,黄页对象的位置信息也同样如此.采用区域方式对用户位置信息进行存储后,对空间数据的访问处理方式又提出了新的挑战.需要提出新的算法解决区域对点(range-to-point),甚至区域对区域(range-to-range)之间的空间查询处理.以下,LaMOC 系统提出了两个新的基于位置服务的应用场景,并提出了与之对应的基于区域的空间数据查询处理算法,为今后新的场景提供了借鉴.

3.3.1 好友推荐与区域对区域 Skyline 算法

移动社交网络应用的一大特点就是帮助用户快速找到与自己位置相近的,并且志同道合的新朋友.考虑以下场景,小明在独自参观上海世博会时,希望在园区内找到一群与他兴趣和年龄等都相似的参观者一起参观.对该场景进行分析后可以发现,小明的位置信息、他的兴趣以及他的年龄等成为了衡量查询结果的几个标准,是一个典型的多目标优化问题.

文献[15]针对以上场景提出了一种基于 Skyline 算法的空间区域对区域算法,目标是从给定的一个 N-维空间的对象集合 S 中选择一个子集,该子集中的对象都不能被 S 中的任意一个其他对象所控制.对象之间的控制关系是指给定一个 N-维空间中的多个对象(对象集 S),若存在这样两个对象 $P=\{p_1,p_2,...,p_n\}$ 和 $Q=\{q_1,q_2,...,q_n\}$,对象 P 在所有维度上的属性值都不比对象 Q 差,且至少在其中一个维度上的属性值优于对象 Q,则称 P 控制 Q.以该场景为例,N-维空间中,用户的位置信息采用了圆形区域表示(以圆心坐标和半径进行描述),即衡量用户之间距离的一个维度是两个圆形区域之间的距离,其他衡量维度如兴趣和年龄等则使用一般数值表示.

图 5 描述了好友推荐场景中的情景.假设参观者小明的位置是图 5(a)中灰色阴影部分,他周围分布了另外 5 位参观者,分别是 C_1~C_5.同时,在图 5(b)图中,在兴趣与年龄两个维度上,采用了相似算法得到了他们与小明的相似度,以数值越小表示在该维度上越接近.在第 1 个比较维度,区域对区域的距离比较中,用户 C_3 控制了其他的 4 位用户,并且在其他两个维度上也能控制 C_2 和 C_5,所以 C_2 和 C_5 因为在各个维度上都差于 C_3,不能进入结果集.C_1 和 C_4 则在兴趣与年龄两个维度上分别不被 C_3 所控制,因此,$\{C_1,C_3,C_4\}$即为最后的查询结果,用户小明得到了系统推荐的这 3 位好友.



(a) Spatial relation of ranges
(a) 空间区域与区域之间关系

	A1	A2
C_1	0.2	0.6
C_2	0.7	0.8
C_3	0.5	0.5
C_4	0.5	0.3
C_5	0.9	0.6

(b) Non-Space attributes of objects
(b) 对象之间的非空间属性

Fig.5 R2R skyline example

图 5 区域对区域 Skyline 例子

文献[15]同时也显示,提出的区域对区域 Skyline 算法在包括距离比较在内的衡量维度数为 4、用户对象数为 4 万时,其算法的执行时间仅为 4s 左右.当衡量维度增加为 6 时,执行时间则增加为 15s 左右.考虑到实际使用中,候选的用户对象数相比实验数据量会减少很多,显然其查询执行时间处于一个能够接受的范围.算法的具体描述可参阅文献[15].

3.3.2 好友集结与区域 Skyline 算法

在日常生活中经常会遇到以下情况:几个好友之间计划下班后聚一聚,如何找到一家相距大家位置都比较

近的饭店?游客准备参观一个城市中的几个景点,如何找到一个距离各个景点位置都比较近的旅馆入住,以方便自己的行程?以上这类场景的显著特点就是包含了多个查询者,且查询目标是向一个被查询对象聚集.本文将该类问题称为多对象聚集问题(multi-object convergence problem,简称 MOC),一个典型的场景即好友集结场景.好友集结场景同样是一个多目标优化问题,最后集结点的衡量标准是通过比较各个可能的集结点与各个查询者之间距离后产生的,每一个集结点与每一个查询者之间的距离就是一个衡量维度,当有 N 位好友希望进行集结推荐时,这即为一个 N-维度的多目标优化问题.

同样,LaMOC 系统为了保护用户的位置隐私,将该问题从点与点之间距离的比较,扩展为点与区域之间的距离比较.用户的位置信息采用满足 k-匿名的一块矩形区域表示,候选集结点的具体位置则为实际点坐标.问题转化以点至各个矩形区域距离为衡量维度的多目标优化问题.在该问题中,首先假设查询者 q 在他所在的矩形区域Ω_q 中的任意位置的出现概率 $P(q)$是相等的,因此被查询对象与查询者之间的距离不是一个确定的值,而是介于 p_i 与 Ω_i 的最短距离 Min$Dist(p_i,\Omega_i)$ 与最大距离 Max$Dist(p_i,\Omega_i)$ 之间(p_i 为某个被查询对象, Ω_i 是某个查询者所在的隐私保护区域),这个距离范围用 $Dist(p_i,\Omega_i)$ 表示.同时,候选集结点的控制关系定义如下:

定义 1. 假设集合 $P=\{p_1,...,p_n\}$ 和集合 $\Omega=\{\Omega_1,...,\Omega_n\}$ 分别是空间中的被查询对象集合和查询者所在的区域集合.现 P 中两点 p 和 p',若对于所有 $\Omega_i\in\Omega$ 都存在 $Dist(p,\Omega_i)\leqslant Dist(p',\Omega_i)$,并且对于某些 $\Omega_j\in\Omega$,存在 $Dist(p,\Omega_j)<Dist(p',\Omega_j)$,则称 p 区域控制 p'.

取某个查询者 q 所在的隐私保护区域Ω_q中任意一点为中心,该点至某个被查询对象 p 的距离为半径,可以做一个圆.这样的圆对 p 可以有无数个,其他被查询对象 $p'\neq p$ 只要出现在任意这样的一个圆中,则 p 无法区域控制 p',因为此时 p'与查询对象 q 的距离小于 p 与其的距离.图 6 中每一个虚线圆表示以查询者 q 在他所在的隐私保护区域中可能出现的一个位置为中心,以这个位置与被查询点 p_1 之间的距离为半径的一个圆,被查询对象 p_3 在 p_1 所对应的某个圆中,所以 p_3 不被 p_1 区域控制.相反,p_2 没有出现在 p_1 所形成的那些园中,所以 p_2 被 p_1 控制.一种较快的衡量两对象之间区域控制关系的方法是,为两个对象做垂直平分线,该线将空间分为两个区域.当查询者所在区域集合都位于其中平分线一边时,在该边的对象则区域控制另一边的对象. $\Omega_i\in\Omega$ 图 6 中,p_1 与 p_2 的区域控制关系也可以用此方法快速得到.最后所有不被其他对象控制的对象集合共同形成集结点的候选集合,即为最后的结果集.

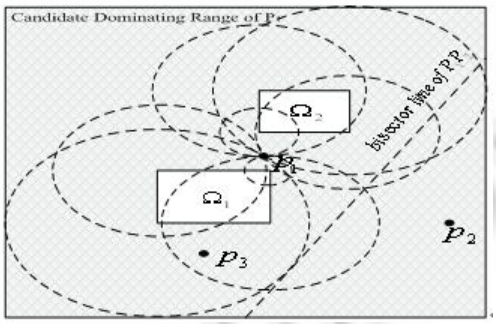


Fig.6 Range-Based spatial skyline query example

图 6 包含 2 个查询区域和 3 个被查询对象的区域 Skyline 查询

由以上判断依据所提出算法的实验结果显示,当集结好友数量分别为 3 人、4 人和 5 人,集合地点候选集的数据量为 1 万时,其算法执行时间分别为 2s、6s 和 7s 左右.在实际使用中,当集结好友数量 N 较大时(N>10),可以就近将几个好友归为一个区域或减少集合地点候选集数量,降低问题的复杂度,以较快速度得到计算结果集.

通过上述对 LaMOC 系统中用户空间信息存储与访问方法的描述,我们可以看到,在位置感知的移动社交网络应用中,对用户位置信息的存储方式可以采用一定算法模糊其精确度,保证该信息即使在遭到恶意攻击时,用户的实际位置也无法快速地被获得.对于经过算法模糊后的位置信息的访问与处理,无论位置信息的表示方式是由点坐标改变为圆形区域还是矩形区域,提出的算法也都能够针对新的应用服务场景,及时有效地对用户

的请求做出回应.

4 结束语

无线通信技术与定位技术的快速发展,以及移动设备的大量普及使得越来越多的用户加入到位置感知的移动社交网络应用使用群体中.一方面,移动社交网络应用促进了人与人之间的沟通交流.另一方面,与他人共享位置信息也使用户面临信息隐私泄露的风险.如何权衡共享位置信息的利与弊,是每一个移动社交网络应用需要解决的问题.

本文首先从对移动社交网络应用中位置隐私问题的四个不同方面分析入手,总结了 3 类在使用该类应用时可能引起用户信息隐私泄露的原因,并提出了 3 种与之相对应的关键保护机制.我们认为,任何移动社交网络应用都必须至少为其用户提供以上 3 种关键保护机制,以保证用户可以活控制自己的位置信息共享方式,降低潜在风险的发生机率.

另外,本文在实验原型系统 LaMOC 中,提出了一种结合了策略与算法的隐私保护机制.基于策略的隐私保护机制为用户提供了更加灵活的隐私保护手段,使其可以针对不同情景,方便快捷地调整隐私策略,在减轻了用户的负担的同时,提高了隐私保护效率.基于算法的隐私保护机制则为用户信息的存储与访问方式提供了新的参考,对于位置信息的不同表现形式,在新的应用场景下,提出的两种算法都能够快速地得到计算结果.希望本文的工作能够起到抛砖引玉的作用,为今后移动社交网络应用的隐私设计提供借鉴.

References:

[1] Dodgeball Inc. http://www.dodgeball.com

[2] Loopt Inc. http://www.loopt.com

[3] Bedo. http://www.bedo.cn

[4] Federal Trade Commission. The Children's Online Privacy Protection Act. http://www.coppa.org/coppa.htm

[5] Directive 95/46/EC. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML

[6] Chang YJ, Liu HH, Chou LD, Chen YW, Shin HY. A general architecture of mobile social network services. In: Yun JN, *et al*., eds. Pro. of the Int'l Conf. on Convergence Information Technology. Gyeongju: IEEE, 2007. 151−156.

[7] Arb M, Bader M, Kuhn M, Wattenhofer R. VENETA: Serverless friend-of-friend detection in mobile social networking. In: Benslimane, *et al*., eds. Proc. of the IEEE WIMOB. Avignon: IEEE, 2008. 184−189.

[8] Markides B, Coetzee M. BlueTrust in a real world. In: Stefan J, *et al*., eds. Proc. of the 3rd Int'l Conf. on Availability, Reliability and Security. Barcelona: IEEE, 2008. 440−445.

[9] Danezis G, Lewis S, Anderson R. How much is location privacy worth? In: Larry G, *et al*., eds. Proc. of the WEIS05. Cambridge: ACM, 2005. 1−13.

[10] Consolvo S, Smith IE, Matthews T, LaMarca A, Tabert J, Powledge P. Location disclosure to social relations: Why, when, & what people want to share. In: Gerrit V, *et al*., eds. Proc. of the SIGCHI Conf. on Human Factors in Computing Systems. Oregon: ACM, 2005. 81−90.

[11] Beach A, Gartrell M, Akkala S, Elston J, Kelley J, Nishimoto K, Ray B, Razgulin S, Sundaresan K, Surendar B, Terada B, Han R. WhozThat? Evolving an ecosystem for context-aware mobile social networks. Journal of IEEE Network, 2008,122(4):50−55.

[12] Lederer S, Mankoff J, Dey AK. Who wants to know what when? privacy preference determinants in ubiquitous computing. In: Gilbert C, *et al*., eds. Proc. of the Conf. on Human Factors in Computing Systems. Fort Lauderdale: ACM, 2003. 724−725.

[13] Prabaker M, Rao J, Fette I, Kelley P, Cranor L, Hong J, Sadeh N. Understanding and capturing people's privacy policies in a mobile social networking application. In: Marc L, *et al*., eds. Proc. of the Workshop on Ubicomp Privacy. Innsbruck: Springer-Verlag, 2007. 1−14.

[14] Chen GL, Rahman F. Analying privacy designs of mobile social networking applications. In: Guo MY, *et al*., eds. Proc. of the IEEE/IFIP Int'l Conf. on Embedded and Ubiquitous Computing. Shanghai: IEEE. 2008. 83−88.

[15] Gu JZ, He L, Yang J, Lu Z. Location aware mobile cooperation-design and system. Int'l Journal of Signal Processing, Image Processing and Pattern Recognition, 2009,2(4):49−60.

[16] Qiao ZZ, Gu JZ, Lin X, Chen J. Privacy-Preserving skyline queries in LBS. In: Tan HH, ed. Proc. of the 2010 Int'l Conf. on Machine Vision and Human-Machine Interface. Kaifeng: IEEE, 2010. 499−504.

[17] Lin X, Li SP, Yang CH. Attacking algorithms against continuous queries in LBS and anonymity measurement. Journal of Software, 2009,20(4):1058−1068 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3428.htm

[18] Pan X, Xiao Z, Meng XF. Location privacy in mobile environment. Journal of Frontiers of Computer Science and Technology, 2007,1(3):268−281 (in Chinese with English abstract).

[19] Barth A, Datta A, Mitchell JC, Nissenbaum H. Privacy and contextual integrity: Framework and Applications. In: Orman H, *et al*., eds. Proc. of the IEEE Symp. on Security and Privacy. Okaland: IEEE. 2006. 183−198.

[20] Ropeik D, Gray G. Risk: A Practical Guide for Deciding What's Really Safe and What's Really Dangerous in the World Around You. New York: Houghton Mifflin, 2002. 17.

[21] Ofcom Inc. http://www.ofcom.org.uk/advice/media_literacy/medlitpub/medlitpubrss/socialnetworking/

[22] Wicaksono N. Connecting windows mobile with vista in new ways. 2007. http://narn.my-sites.net

[23] Zhang ZY, Yang J, Gu JZ, Xu YC. Event based semantic location model in cooperative mobile computing. In: Pan JS, *et al*., eds. Proc. of the 9th Int'l Conf. on Hybrid Intelligent Systems. Shenyang: IEEE, 2009. 203−208.

[24] Ma J, Lu H, Guo YF. Region based cameraphone image retrieval incorporated with metadata. Journal of Image and Graphics, 2007, 12(10):1766−1769 (in Chinese with English abstract).

[25] Zhang L, Wang JY, Dai N. Stimulation and study on the signal acquisition and tracking of GPS. Journal of Modern Defense Technology, 2008,36(2):107−110 (in Chinese with English abstract).

[26] Yi J, Zhang L, Wang JY, Shu R. The location services analyze and application research based on satellite positioning. Journal of World SCI-TECH R&D, 2008,30(3):328−330 (in Chinese with English abstract).

[27] Xu YC, Gu JZ, Yang J, Zhang ZY. An ontology-based approach for mobile personalized recommendation. In: Pan JS, *et al*., eds. Proc. of the 9th Int'l Conf. on Hybrid Intelligent Systems. Shenyang: IEEE, 2009. 336−339.

[28] Gedik B, Liu L. Protecting location privacy with personalized *k*-anonymity: Architecture and algorithms. IEEE Trans. on Mobile Computing, 2008,7(1):1−18.

附中文参考文献:

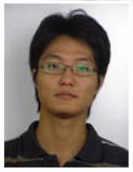
[17] 林欣,李善平,杨朝晖. LBS 中连续查询攻击算法及匿名性度量.软件学报,2009,20(4):1058−1068. http://www.jos.org.cn/1000-9825/3428.htm

[18] 潘晓,肖珍,孟小峰.移动环境下的位置隐私.计算机科学与探索,2007,1(3):268−281.

[24] 马桔,路红,郭跃飞.融合相机元信息的基于区域的手机图片搜索.中国图像图形学报,2007,12(10):1766−1769.

[25] 张雷,王建宇,戴宁.GPS 信号捕获跟踪的仿真分析与研究.现代防御技术,2008,36(2):107−110.

[26] 易炯,张雷,王建宇,舒嵘.基于卫星定位的位置服务分析及其应用研究.世界科技研究与发展,2008,7(1):1−18.



谈嵘(1984－),男,上海人,博士,主要研究领域为情景感知计算,隐私保护.

林欣(1981－),男,博士,主要研究领域为隐私保护,时空数据库,情景感知计算.

顾君忠(1949－),男,教授,博士生导师,主要研究为情景感知计算,分布式数据库管理,多媒体信息处理.

陈鹏(1985－),男,博士,主要研究领域为数据库管理,不确定性数据.

杨静(1976－),女,副教授,主要研究领域为语义信息处理,智能多媒体.

乔哲峰(1986－),男,硕士,主要研究领域为隐私保护,时空数据库.