目录

**移动商务的用户隐私保护**

摘要

移动商务的隐私泄露问题已经在全世界引起越来越多的关注。虽然有许多与信息隐私和隐私保护技术相关的论文，但很少是基于特定的移动商务模型来研究匿名模型和隐私保护算法的。一种基于位置服务的移动商务的隐私保护服务框架已经建立。根据移动用户对于隐私信息的个性化定义，（K，L，P）匿名模型正式提出。基于该匿名模型，提出了用于匿名组交换和合并进程的新的隐私保护算法（EMAGAS），包括设置一个最小初始K-匿名集，交换和合并的过程被形式化的描述。EMAGAS可以用来保护用户的位置，标识出公路网络上移动用户的位置和其他敏感信息。EMAGAS的可用性是通过一个例子来说明。最后，基于移动用户在真实的公路网上的隐私分布，可以通过实验来验证可行性和EMAGAS的优势。

关键词：匿名模型；个性化隐私信息；隐私保护算法；服务框架；用户隐私保护

Abstract

The risk of privacy disclosure in mobile commerce has received increasing attention worldwide. Although many papers related to information privacy and privacy-preserving technologies exist, few are based on a particular mobile commerce model to study the anonymity models and privacy-preserving algorithms. A privacy-preserving service framework for the mobile commerce alliance providing location-based services is established. According to the defined personalized privacy profile of the mobile user, a (K, L, P)-anonymity model is formally described. Based on the model, a new privacy-preserving algorithm for exchanging and merging processes for generating anonymity sets (EMAGAS) is proposed, which features the construction of minimal initial K-anonymity sets, an exchanging process and a merging process. The processes of exchanging and merging are formally described. EMAGAS can be used to protect the location, identifier and other sensitive information of the mobile user on a road network. The availability of EMAGAS is illustrated by an example. Finally, based on a real road network and generated privacy profiles of mobile users, the feasibility and advantages of EMAGAS are experimentally validated.

Keywords: Anonymity model; Personalized privacy profile; Privacy-preserving algorithm; Service framework; User privacy protection

1. 介绍

移动电子商务是指使用移动手持设备，如蜂窝电话，并通过移动网络的个人数字助理（PDA）所进行的电子商务活动。与传统电子商务相比，移动电子商务有一些新功能，包括移动性，即时性，个性化和方便性。基于位置的服务（LBS），是一种移动用户访问的信息服务，它是使用基于移动通信的技术，例如全球定位系统（GPS），无线局域网（WLAN）和移动设备的地理位置信息蜂窝网络。近年来，随着新的应用信息和通信技术的普及，移动电子商务得到迅速的发展。其中使用最广泛的是基于位置的移动商务应用­—移动广告。使用新型移动商务应用向用户提供位置服务已经成为流行，同时有利于资源共享的各种类型的移动电子商务联盟正在出现。

使用LBS，移动用户通常需要向服务提供商发送其查询请求和精确的位置。服务提供商可能会收集，处理和存储用户的地理位置信息，所以与地理位置相关的隐私问题也引起了越来越多的关注。虽然有很多与隐私保护技术有关的论文，但是很少是基于特定的移动商务模型来研究的匿名模型和隐私保护算法。本研究试图解决以下问题：(1)为移动电子商务提供个性化的隐私保护服务框架。(2)如何在移动电子商务的背景下定义个性化的移动用户的隐私要求。(2)根据定义的匿名模式，建立一个新的隐私保护算法。

在第2节，在了解与移动电子商务中的隐私信息相关的概念和工作后，对在移动环境中常用的隐私保护技术进行了讨论。在第3节，建立了为移动电子商务提供位置服务的隐私保护服务框架。在第4节，根据移动用户定义的个性化的隐私信息建立了（K,L,P）-匿名模型。基于匿名模型，一种在生成匿名集（图像）中基于交换和合并过程的新的隐私保护算法被提出。详细讨论了用户的交换和合并过程。在第5节，通过一个例子来说明EMAGAS的可用性。在第6节，基于一个真实的路网和移动用户的隐私信息，可行性和EMAGAS的优势通过实验验证。结论最后介绍。

1. 相关工作
   1. 移动商务中的隐私问题

移动电子商务应用的隐私问题主要涉及移动用户的个人基本信息，移动设备信息，手机支付信息，位置信息，社会关系信息等隐私相关的信息。从本地传感器，如GPS接收器，照相机，麦克风和加速度计等可以由第三方应用程序使用，其中的信息可以看作是智能手机上的用户隐私敏感数据。数据隐私和安全是移动电子商务主要问题。位置相关的隐私信息的泄露是影响移动商务发展的一个关键因素，所以在建立研究模型时，要考虑到在移动商务中提供的隐私信息中可以影响行为意向的因素。

位置信息是一种特殊类型的隐私信息。在移动环境下，用户身份的隐私是指当攻击者获取并分析用户的其他身份相关的信息时，防止从公开内容中查询用户的身份标识。例如，位置信息可以在查询中被用作一个伪标识。在提供LBS功能的移动应用程序中，敏感数据可以是一般的相关信息，如健康信息可作为服务请求的一部分来发送财务数据，它也可能是由服务提供者收集移动用户的时空信息，例如，用户在特定时间的位置信息，运动模式或个人利益点（经常访问特定的商店，俱乐部或机构）。

用户担心服务提供商会共享收集，储存和使用他们的私人信息，认为他们的位置信息的泄露可能会为他们带来风险。然而，要使用LBS，移动用户必须通过因特网发送其位置信息给服务提供商，这可能导致两种类型的隐私问题。一种是攻击者可能通过定位位置传输设备或窃听的数据传输信道获取原始位置数据，从多个来源获取的数据中收集与位置有关的隐私信息。另一种是，服务提供商可以通过分析查询请求的信息和准确的位置数据来收集用户的个人敏感信息，甚至转售或为了利润泄露隐私信息。因此，服务提供商应该采取措施，以减少用户的隐私顾虑，来促进移动商务的发展。

* 1. 移动环境下的隐私保护技术

**2.2.1 基于公路网上的隐私保护技术**

研究位置隐私保护技术可分为基于自由空间和基于受约束空间两类。对于基于自由空间技术中，假定该用户可以在欧几里得空间中自由移动而不约束。然而，在实际的移动环境中，无论他们要求什么类型的LBS，移动用户通常在约束网络中移动，如公路网和铁路网。采用基于欧几里得空间道路网络环境的位置匿名的方法可能导致泄露隐私。

在过去的几年中，以保护道路网络中的移动用户的位置隐私为目的，各种解决方案已经被提出。大多数现有的研究结果集中在基于所述位置的K-匿名模型来保护用户位置和身份识别码的信息。K-匿名模型即将用户的精确位置用一个至少包含其它k- 1个人的空间区域来代替,从而使攻击者无法将某一个人的位置信息通过推理攻击的方式与其身份相匹配。

Mouratidis首次应用的位置K-匿名模型公路网。公路网络建模为无向图。假设图1中用户A，B和C构成匿名集，它遵循位置3—匿名模式。匿名服务器发送匿名集来为平台提供LBS。因此，对于攻击者来说很难区分匿名集中的三个用户。Smith提出了关于在受限网络移动模型中保护移动用户位置隐私的问题的研究，并提出了一个一般匿名位置模型，XSTAR，用于通过公路网络的隐私保护服务。Chow提出了一个适用于公路的网络环境的位置匿名算法，其中一个用户位置在一组最小总长度为L的连接道路段中被匿名，路段中至少包括K个用户，并考虑在地理环境下附近的道路段。所有上面提到都是在保护移动用户的位置和身份标识，没有考虑信息披露的敏感性研究工作。

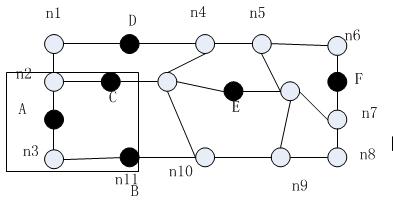


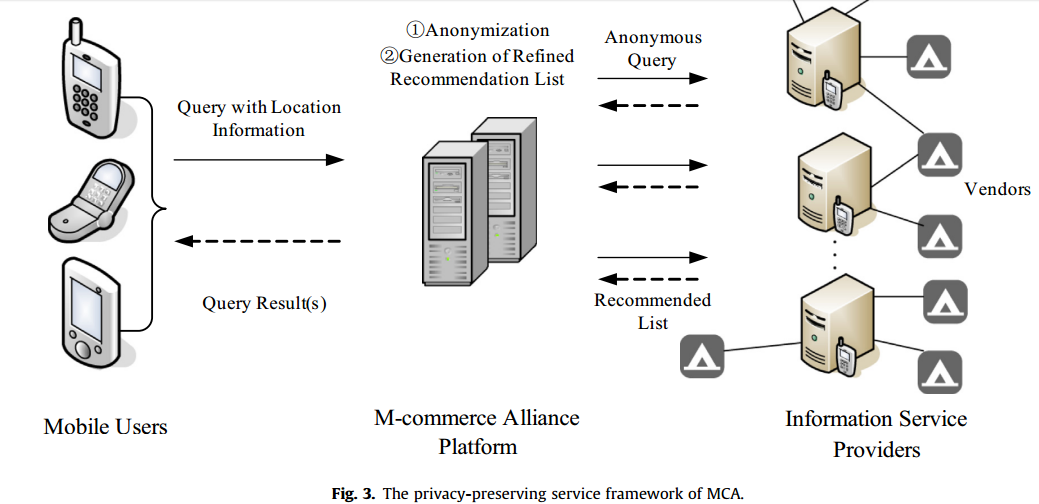
图1 公路网建模图

已经有一些关于移动电子商务的敏感信息保护的研究。有一个是TaintDroid，通过第三方应用程序的流量跟踪隐私敏感数据并扩展到Android手机平台。Pan是第一个研究公路网络中的用户敏感信息的保护技术。他们提出了（K，L，P）-匿名模型和P3RN隐形算法。这种方法支持个性化的隐私要求，确保用户的位置信息安全。根据改进该方法的主要方法，我们将提出一个新的隐私保护算法。

1. 移动电子商务的隐私保护服务框架
   1. MCA的隐私保护服务框架

随着移动电子商务的快速发展，最近几年出现了各种类型的移动电子商务联盟。与重点是B2B营销的商业网络（IBN）相反，我们探讨的是向移动电子商务联盟（MCA）提供B2B2C的服务。电子商务联盟旨在提供可信的，可靠的和增值的IT基础设施服务，促进资源共享，促进参与玩家之间的双赢关系。在本文中，我们认为MCA作为一个值得信赖的信息服务平台，它与许多移动信息服务提供商合作，每一个都拥有自己的特定供应商的资源。

MCA的隐私保护服务框架如图3所示。它有四类实体组成：MCA平台，信息服务供应商，拥有实体店的供应商，以及移动用户。各类实体有它自己的权利和义务。MCA平台作为一个可信的提供匿名服务和全面的基于位置的信息服务的第三方，包括根据移动用户的查询请求和目前的位置的供应商推荐。移动用户有权力提出隐私权要求。



* 1. 基于MCA供应商推荐服务模式的隐私保护

服务模型如下:

* 移动用户在MCA平台上注册他们的基本信息，这些信息可以被访问。
* 当接收到查询请求时，MAC平台将执行一个基于用户位置查询的匿名算法， 产生一个可以满足指定匿名模式要求的匿名集。
* 向MCA的信息服务提供商发送匿名查询请求和匿名集。
* 各信息服务提供商使用特定的算法来处理查询，并将查询结果返回到MCA平台，如推荐厂商的名称，主页链接和其他需要的信息。
* 基于查询用户的精确位置和多个信息服务提供商的推荐供应商列表，在MCA平台上使用排列算法并发送最佳推荐列表给查询用户。
* 当用户采用该建议，它可以被看作是一种有效的建议，或有效的反馈。为了提高推荐质量，所收集的反馈可以作为推荐算法排名的指标之一。

根据上述的供应商推荐模型，MCA平台接收用于查询用户精确的位置信息，并生成匿名查询，而各信息服务提供商必须基于匿名集执行它自己的推荐算法和生成其向服务商提供的推荐列表。我们会根据不断的交换和合并过程为平台建立一个匿名模型并给出一个隐私保护算法。

1. 公路网络的个性化隐私保护算法
   1. 基本定义

信息敏感性指对可能会泄露的信息的访问的控制，如果泄露或与他人共享可能导致信息失去安全性。不同的移动用户可以基于不同的隐私倾向分配给相同类型的信息以不同的灵敏度水平。移动用户的信息隐私需求取决于其信息敏感性和个人性格。

参照Pan等人的工作，提交给匿名服务器的所有查询请求可以被划分为不同的灵敏度类别，以及每个类别都标有一个[0，1]之间的值来表示相应的信息灵敏度。标记值越大，敏感度级别越高。例如，查询Q1，找到最近的医院，灵敏度水平标记为0.75;另一个查询Q2，找到最近的小学，灵敏度水平标记为0。因此，Q1比Q2更敏感。

我们首次正式定义移动用户个性化的隐私信息。

**定义1**（个性化隐私信息）鉴于该移动用户可以指定一个人的个人隐私信息，隐私信息可以被描述为四元组（k,qsr,sd,p）。

在此定义中，匿名需求（k）代表移动用户可以接受的最小匿名级别，这意味匿名集中至少有K个移动用户。k的值越大，移动用户识别符的匿名程度越高。查询灵敏度需求（qsr）表示移动用户允许最大的查询灵敏度。如果查询Q的灵敏度比qsr大，则Q对于用户来说是敏感查询;否则，Q是对于用户是一个非敏感的查询。qsr的值越大，敏感查询的耐受性越高。段多样性要求（sd）代表匿名地区不同路段的最小数量。sd的值越大，移动用户的位置匿名的水平越高。设定灵敏度需求（p）的意味着匿名集中移动用户对敏感查询的最大容限比。p的值越大，设定的灵敏度越严格。

例如，一个移动用户u的隐私信息为（4，0.25，3，0.5）。这意味着，u要求匿名集包含至少4个不同的移动用户和匿名区域至少有3个不同的道路段。如果一个查询的敏感度大于0.25，U认为这个查询是一个敏感的查询，要求所涉及的匿名敏感查询的最大容限比设定不大于0.5。

**定义2（匿名集）**这是由在匿名区域的移动用户组成。对于用户的匿名集设为AS，| AS |代表在AS的移动用户的数目。

对于所有的，u的匿名需求是u.k，其查询灵敏度需求是u.qsr，其集灵敏度需求是u.p。u发送的查询灵敏度是u.qs。另外，在AS中的用户的最大匿名需求被表示为最大u.k，匿名集中道路段的数量被表示为AS.Count\_S，u的敏感查询的数量是AS.Count\_SQu。

**定义3 （匿名算法）**如果匿名集满足三个条件，则称为是（K，L，P）模型。条件为：（1）K-位置匿名，，其中，K= Max u.k;（2）路段多样性，，其中表示匿名域的路段数，，有L = Max u.sd；（3）P-匿名敏感信息，例如 u.p AS. Count\_SQu/.

通过应用（K，L，P）-anonymity模型，用户在移动环境中的隐私信息可以被全面保护。第一个条件意味着AS中的用户具有K共享属性，这可以防止用户的标识符和位置被公开查询。第二个条件保证已公布的路段多样性和用户确切的位置查询对于攻击者是隐藏的。最后，第三个条件有助于保护敏感信息。

例如，令AS1 = {u1，u2，u3}和AS2 = {u4，u5，u6，u7}。 移动用户的个性化隐私信息如表1a所示，查询和标签敏感度值如表1b所示。根据（K，L，P） - 匿名模型，值的所有变量列于表1c。可以看出u6不满足（K，L，P） - 匿名模型。 可以表明，AS1满足（K，L，P） - 匿名模型，而AS2不满足。

表1a个人隐私信息

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mobile user | K | qsr | sd | p |
| u1 | 2 | 1 | 2 | 0.4 |
| u2 | 3 | 0.6 | 2 | 0.5 |
| u3 | 3 | 0.4 | 2 | 0.8 |
| u4 | 3 | 0.6 | 2 | 0.6 |
| u5 | 2 | 0.4 | 2 | 0.8 |
| u6 | 2 | 0.25 | 2 | 0.5 |
| u7 | 2 | 0.5 | 2 | 1 |

表1b Query and query sensitivity.

|  |  |  |  |
| --- | --- | --- | --- |
| Anonymity set | Mobile user | Query | Query sensivity |
| AS1 | u1 | Q1 | 0.25 |
| u2 | Q2 | 0.5 |
| u3 | Q3 | 1 |
| AS2 | u4 | Q4 | 0.5 |
| u5 | Q5 | 1 |
| u6 | Q6 | 0.5 |
| u7 | Q7 | 0 |

表1c Values of the variables

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Anonymity.set | Mobile user | ｜AS｜ | Max(u.k) | AS.Count\_S | Max(u.sd) | AS.Count\_SQu | u.p | AS.Count\_SQu/｜AS｜ |
| AS1 | u1 | 3 | 3 | 3 | 2 | 0 | 0.4 | 0 |
| u2 | 1 | 0.5 | 1/3 |
| u3 | 2 | 0.8 | 2/3 |
| AS2 | u4 | 4 | 3 | 4 | 2 | 1 | 0.6 | 1/4 |
| u5 | 3 | 0.8 | 3/4 |
| u6 | 3 | 0.5 | 3/4 |
| u7 | 1 | 1 | 1/4 |

* 1. 匿名算法 EMAGAS

为了保护移动用户的隐私信息，MCA平台必须将移动用户向平台发送的所有查询请求分组为满足（K，L，P） - 匿名模型的不同匿名集。我们提出了一种通过交换和合并用于生成匿名集的过程的算法(EMAGAS).EMAGAS的特征包括：

（1）构建满足k-匿名的最小初始匿名集;

（2）对于不满足（K，L，P）- 匿名模型的初始匿名集，交换相邻匿名集中的用户，然后合并相邻的匿名集;

（3）对于在交换和合并处理完成之后不满足（K，L，P） - 匿名模型的匿名集，添加虚拟对象使它们满足（K，L，P） - 匿名模型。 EMAGAS算法的流程图如图4所示。

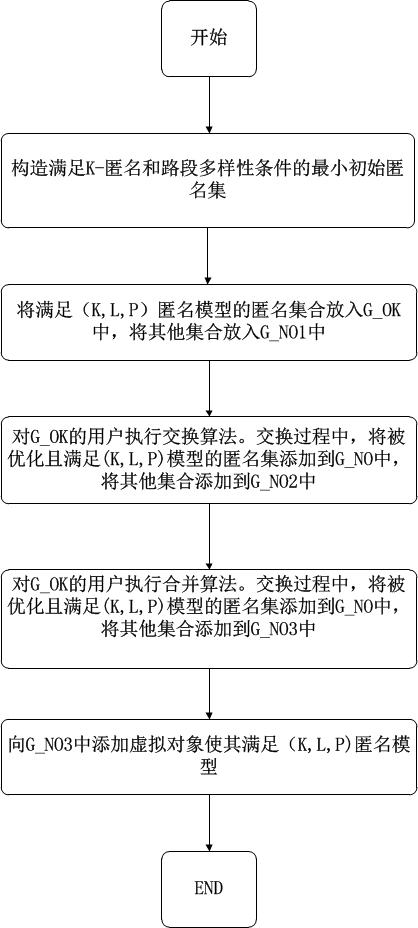


图4 EMAGAS算法的流程图

图4中的步骤如下：

（1）使用深度优先搜索算法（DFS）遍历道路网络。 在遍历中，路段和移动用户按顺序编号。

（2）选择道路段上的移动用户作为新匿名集的第一个组成部分，该移动用户未包含在任何构造的匿名集中。查找并按顺序向新匿名集中添加尚未包含在任何构造的匿名集中的相邻用户，直到该集合中的用户数量满足用户的最大匿名性需求，并且满足路段多样性的条件。

（3）重复步骤2，直到每个移动用户都包含在构建的匿名集中。

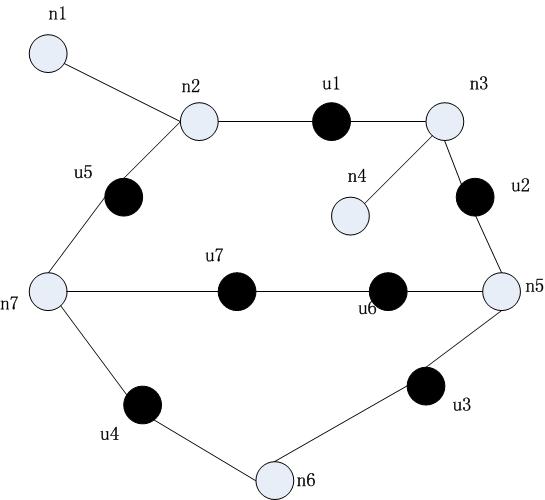
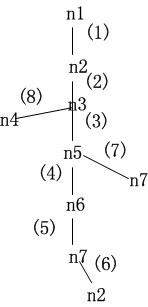
 

图5（a） 图5（b）

道路网络示例如图5（a）所示，对a的DFS遍历结果如 图5（b）所示。用户{u1，u2，u3，u4，u5，u6，u7}的匿名性需求是{2,3,4,4,3,2，2}，并且路段需求{2,2,2,2,2,2,2}。 使用上述方法，建立两个初始匿名集{u1，u2，u3，u4}和{u5，u6，u7}。

* + 1. 用户交换算法

不满足（K，L，P） - 匿名模型的初始匿名集使用用于交换用户的算法来处理。 处理步骤如下所述。

（1）按照不满足（K，L，P）匿名模型的不安全用户的数量，按照升序对G\_NO1中的初始匿名集进行排序。

（2）从G\_NO1获取未处理的匿名集AS1。 将G\_NO1中未处理且与AS1相邻的匿名集提取到G\_Temp，并降序排列它们与AS1之间的邻接度。

（3）从AS1获取未处理的不安全用户u1。

（4）依次从G\_Temp获取匿名集AS2。

（5）从AS2获取未处理的不安全用户u2。 如果这四个条件成立：（a）u2.k  | AS1 |;（b）u2.sd  AS1.Count\_S，如果u1和u2交换AS1.Count\_S的值不会变小;（c）u2.qs  u1.qs; （d）u2.p i / K，其中i表示由AS1 中的用户发送的对u2的敏感查询的数量。

（6）如果AS1满足（K，L，P） - 匿名模型，将AS1移动到G\_OK;

重复步骤（3）到（6）通过在不同的匿名集中交换不安全的用户以优化匿名集。

（7）将G\_NO1中的剩余匿名集置于G NO2中。

在交换用户的过程中，步骤5中的四个条件确保AS2中的不安全用户u2成为AS1中的安全用户，并且AS1中的现有安全用户在交换操作之后仍然安全。也就是说，AS1中的安全用户不会因交换u1和u2而受到负面影响。如果AS1包含多个不安全用户，则按不安全用户按照其发送的查询的敏感度降序排序，然后获取当前未处理的不安全的用户。

|  |
| --- |
| 输入：G\_NO1,G\_OK,privacy profiles,query sensitivies  输出：G\_NO2,G\_OK  1: 按照不安全用户的数量对G\_NO1中的匿名集进行排序  2: 将G\_NO1中的所有匿名集标作未处理  3: If G\_NO1中有未处理匿名集　then  4: 取G\_NO1中未处理的匿名集AS1  5: 将AS1标记为已处理  6: 在G\_temp中排列与AS1相邻的未处理匿名集  7: for each user u1 in AS1  8: if G\_temp{} then  9: 取G\_temp 中的匿名集AS2  10: 从AS2中选择不安全用户添加到UU  11: if UU{} then  12: 从UU中选择用户u2  13: m=AS1.Count\_S //如果u1和u2交换  14： i=AS.Count\_SQu2 //如果u1和u2交换  15： if u2.k| AS1 | and u2.sdAS1.Count\_S  16: and AS1.Count\_Sm and u2.pi/k  17: then 交换u1和u2  18: 将u2标记为安全用户  19： if AS1满足（K,L,P）匿名模型  20： then 将AS1添加到G\_OK  21： go to Step 3  22: go to Step 11  23: else go to Step 8  24: end of for statement  25: go to Step 3  26: 将G\_NO1中剩余匿名集添加到G\_NO2 |

* + 1. 用户合并算法

通过执行用户交换算法，一些初始匿名集可以满足（K,L,P）模型。G\_NO2中剩余的匿名集可以通过合并不同匿名集中的用户来进一步的优化。首先，给出p-关键用户和p-不满意用户的定义。

**定义4**（p-关键用户）对于匿名集AS中的用户u，向AS中合并一个用户u‘（u.qs）后，u不再满足集敏感度需求，则称u为AS中的一个p-关键用户。

**定义5**（p-不满意用户）如果匿名集AS中的用户user不满足集敏感度需求，则这个用户为p-不满意用户。

例如，AS={u1,u2,u3}。如果AS.Count\_SQu1=1，u1.p=0.4，AS.Count\_SQu1/｜AS｜=1/3。这就意味着u1满足集敏感度需求。相反如果AS.Count\_SQu2=2，u2.p=0.6，AS.Count\_SQu2/｜AS｜=2/3，这就意味着u2不满足集敏感度，所以u2是p-不满意用户。当用户u’（u’.qsu1.qsr）加入到AS，AS.Count\_SQu1/｜AS｜=2/4，则u1不再满足集敏感度需求，所以u1是p-关键用户。

用户的合并过程如下：

1. 将G\_NO2中的匿名集按其不安全用户的数量升序排列。
2. 将G\_NO2中的所有匿名集标作未处理。
3. 从G\_NO2获取一个未处理的匿名集AS1，提取未处理且与AS1相邻的匿名集添加到G\_Temp，按照它们和AS1之间的邻接度进行降序排列。
4. 从G\_Temp依次获得匿名集AS2。
5. 提取AS2中的不安全用户到UU。所选择的用户发送的查询请求的敏感度要小于AS1中p-关键用户和p-不满意用户的最小查询敏感度需求。
6. 从UU中取一个用户，将它合并到AS1。
7. 如果AS1满足（K,L,P）匿名模型，则将AS1添加到G\_OK。
8. 重复步骤（3）到步骤（7），通过合并不安全用户来优化匿名集。
9. 将G\_NO2的匿名集添加到G\_NO3。

|  |
| --- |
| 输入：G\_NO1,G\_OK,privacy profiles,query sensitivies  输出：G\_NO3,G\_OK  1: 按照不安全用户的数量对G\_NO2中的匿名集进行排序  2: 将G\_NO2中的所有匿名集标作未处理  3: If G\_NO2中有未处理匿名集　then  4: 取G\_NO2中未处理的匿名集AS1  5: 将AS1标记为已处理  6: 在G\_temp中排列与AS1相邻的未处理匿名集  7: if G\_temp{} then  8: 取G\_temp 中的匿名集AS2  9: m1=minimum u.qsr // u为p-关键用户  10： m2= minimum u.qsr // u为p-不满意用户  11: min\_qsr=min(min1,min2)  12: for each user u in AS2  13: if u.qsmin\_qsr  15： then 将u添加到UU  16： 从UU中选择一个用户u  17: 将u添加到AS1  18： if AS1满足（K,L,P）匿名模型  19： then 将AS1添加到G\_OK  20: go to Step 3  21: else go to Step 15  22: go to Step 7  23: 将G\_NO2中剩余匿名集添加到G\_NO3 |

1. 在图像中应用EMAGS

为了说明所提出的算法EMAGAS的能力，让我们假设， 图6是道路网络示例。 使用我们提供的方法，生成初始匿名集。个人隐私资料和18个路网上的用户的查询敏感性，以及生成的初始匿名集，如表2所示。

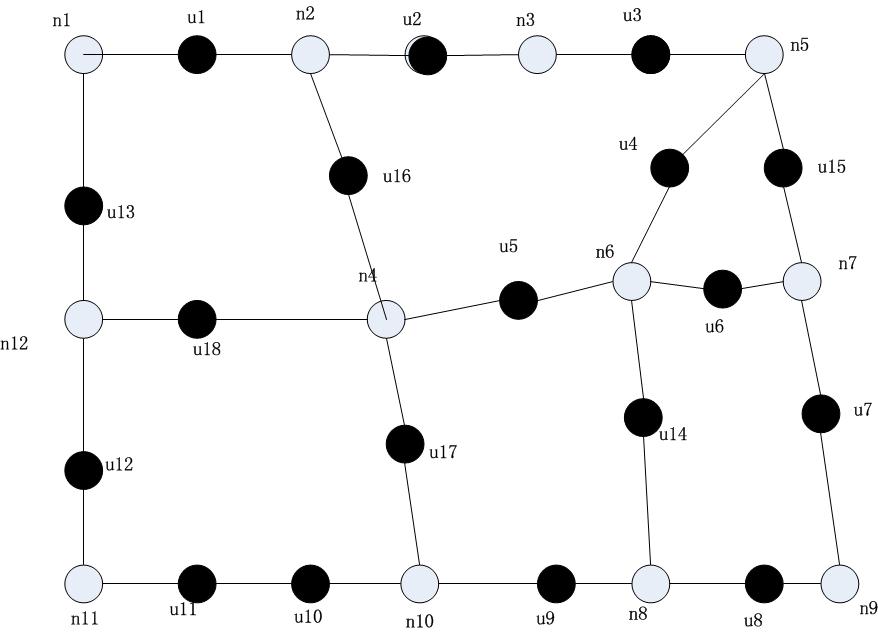


图6 道路网络示例

表2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anonymity sets | Users | k | sd | qsr | p | qs |
| AS1 | u1 | 2 | 2 | 0.5 | 0.6 | 0.5 |
| u2 | 2 | 2 | 0.4 | 0.6 | 1 |
| AS2 | u3 | 3 | 2 | 0.4 | 0.7 | 0 |
| u4 | 3 | 2 | 1 | 0.4 | 0.25 |
| u5 | 3 | 2 | 0.6 | 0.5 | 0.5 |
| AS3 | u6 | 2 | 2 | 0.4 | 0.8 | 1 |
| u7 | 2 | 2 | 0.25 | 0.4 | 0.5 |
| AS4 | u8 | 3 | 2 | 0.5 | 1 | 0 |
| u9 | 3 | 2 | 0.5 | 0.7 | 0.5 |
| u10 | 2 | 2 | 0.6 | 0.5 | 0.75 |
| AS5 | u11 | 3 | 2 | 0.4 | 0.7 | 0.25 |
| u12 | 2 | 2 | 0.7 | 0.4 | 0.5 |
| u13 | 2 | 2 | 0.5 | 0.5 | 0 |
| AS6 | u14 | 3 | 2 | 0.3 | 0.7 | 1 |
| u15 | 2 | 2 | 0.8 | 0.3 | 0.25 |
| u16 | 3 | 2 | 1 | 0.3 | 0.25 |
| AS7 | u17 | 3 | 2 | 0.5 | 0.7 | 1 |
| u18 | 2 | 2 | 0.6 | 0.8 | 0.75 |

表3 Anonymity sets in G\_NO2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anonymity | users | k | sd | qsr | p | qs |
| AS3 | u6 | 2 | 2 | 0.4 | 0.8 | 1 |
| u7 | 2 | 2 | 0.25 | 0.4 | 0.5 |
| AS6 | u14 | 3 | 2 | 0.3 | 0.7 | 1 |
| u2 | 2 | 2 | 0.4 | 0.6 | 1 |
| u16 | 3 | 2 | 1 | 0.3 | 0.25 |
| AS7 | u17 | 3 | 2 | 0.5 | 0.7 | 1 |
| u18 | 2 | 2 | 0.6 | 0.8 | 0.75 |

表4 Anonymity sets in G\_NO3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anonymity sets | Users | k | sd | qsr | p | qs |
| AS3 | u6 | 2 | 2 | 0.4 | 0.8 | 1 |
| u7 | 2 | 2 | 0.25 | 0.4 | 0.5 |
| AS6 | u14 | 3 | 2 | 0.3 | 0.7 | 1 |
| u2 | 2 | 2 | 0.4 | 0.6 | 1 |

根据定义3的描述，可以看出AS1，AS3，AS6和AS7不满足（K，L，P） - 匿名模型。匿名集AS1，AS3，AS6和AS7使用用于交换用户的算法来处理。交换AS1中的用户u2和AS6中的用户u15，因此AS1满足（K，L，P） - 匿名模型。 不同匿名集中的其他不安全用户对不能交换。如表3所示，将不满足（K，L，P） - 匿名模型的剩余匿名集放入G NO2。通过应用合并用户的算法，AS6中的用户u16被合并到AS7中，然后AS7满足（K，L，P） - 匿名模型。 如表4所示，不满足（K，L，P） - 匿名模型的剩余匿名集被放入G\_NO3。最后，我们可以通过在其中插入虚拟对象，使G\_NO3中的匿名集合满足（K，L，P） - 匿名模型。

1. EMAGAS算法的性能实验分析

基于真实世界道路网络通过实验比较了EMAGAS算法与P3RN的性能（Pan等人，2014），对于先前讨论的用户的隐私信息采用三个度量：平均虚拟比率，查询成本和平均匿名时间。现有的在基于道路网络位置隐私保护工程中，P3RN似乎是解决查询敏感性问题的首选算法。我们实现的两种算法都是在JAVA中实现的。我们也在台式电脑上运行它们一个双核AMD 2.2 GHz处理器和4 GB的主存储器。

* 1. 平均虚拟比率

可以通过向不满足模型的匿名集添加虚拟对象来增加匿名集的成功率。虚拟比率是指匿名集中添加的虚拟对象的百分比。 使用EMAGAS算法，通过在交换和合并处理完成之后添加虚拟对象，可以实现100％成功生成满足（K，L，P） - 匿名模型的匿名集。使用D =来计算平均虚拟比率，其中num表示虚拟对象的数量，usernum表示用户的总数。

实验结果如图8a所示. 当使用虚拟比率来比较两种算法时，EMAGAS的性能比P3RN差，但是随着maxk的值的增加，性能差距变小。将虚拟对象添加到为P3RN设置的匿名集中是因为匿名集不满足P条件，而对于EMAGAS来说可能是不满足条件（K，L，P）中的任一个，因此需要添加更多的虚拟对象。然而，与P3RN相比，EMAGAS可以生成更小的匿名集和匿名域，这有助于基于匿名查询信息为LBS提供商提供更准确和有用的推荐。

* 1. 查询成本

处理基于位置的查询的成本包括LBS平台的查询执行成本和将查询结果传送回移动用户的通信成本（Wang和Liu，2009）。 在本研究中，我们只考虑查询执行成本。给定对应于匿名集AS的路段SS的集合，VS是SS中的路段的端点集合。对于所有如果存在一个端点为v并且另一个端点不在VS中的道路段，则v被称为开放端点。让VS中的开放端点集合为OVS。 然后计算平均查询成本如下：（1）通过QCost（AS）= AS.Count\_S + | OVS |计算每个匿名性集合的查询成本，其中AS.Count\_S表示SS中的路段数量，| OVS | 表示设置OVS中开放端点的数量。然后计算平均查询成本与成本¼开i¼1ðASi：计数，其中n是匿名的数目集。

查询成本的实验结果如图 8b所示。 两种算法的平均查询成本随着maxk的增加而增加，当maxk 10时，EMAGAS的平均查询成本低于P3RN的平均查询成本。因此，当比较两种算法的查询成本时，EMAGAS的性能优于P3RN的性能。 原因在于在EMAGAS算法中生成匿名集的过程中的改进。使用P3RN，每个初始匿名集的大小是maxk，而由合并操作产生的匿名集的大小可以变为2 maxk。使用EMAGAS算法，初始匿名集的大小被动态地确定为不大于maxk。 合并操作涉及将来自相邻匿名集的一个用户带到指定的匿名集中。实验结果分析表明，使用EMAGAS生成的匿名集的平均大小小于使用P3RN生成的匿名集的平均大小。这是为什么EMAGAS的查询成本低于P3RN的主要原因。

* 1. 平均匿名时间

平均匿名时间表示算法的效率。 时间越短，算法的效率越高。 平均匿名化时间由下式给出： ，其中表示道路上的用户总数网络，T表示匿名化算法的运行时间。 图8c示出了两种匿名算法的匿名时间。 当maxk等于或大于10时，使用P3RN的匿名化时间约为0.8ms，而使用EMAGAS的匿名化时间变为小于0.1ms。因此，IMAGES的效率远远高于P3RN的效率。 主要原因在于两种算法的主要步骤中的不同处理方法。

1. 结论
   1. 结果讨论

在本文中，为提供LBS的商业联盟（MCA）建立了一个隐私保护服务框架，这使得多个信息服务提供商的服务资源得到整合，有助于实现联盟中的所有参与者的双赢。可以防止移动用户的隐私信息被信息服务提供商和供应商收集和滥用，并且用户可以接收全面的信息服务。

实现MCA的建议的隐私保护服务框架的关键之一是基于适当的匿名模型的匿名化过程。个性化隐私信息被正式定义为四元组（k，qsr，sd，p）并可用于描述移动用户的隐私要求，包括匿名性，查询灵敏度，段多样性和设置敏感性要求。 根据定义的个性化隐私信息，引入并正式描述（K，L，P） - 匿名模型。

基于（K，L，P） - 匿名模型，提出了一种名为EMAGAS的新的隐私保护算法，其特征在于构造最小的初始K-匿名集，交换过程和合并过程。交换和合并的过程将逐步详细讨论和形式化的描述。 所提出的EMAGAS算法用于保护移动用户的位置和标识符，但也有助于保护敏感信息。为了证明EMAGAS的可行性，我们通过示例应用程序说明了每个主要步骤的结果。 使用真实的道路网络和移动用户的隐私信息，通过使用虚拟比率，查询成本和匿名化时间的度量实验分析和比较EMAGAS和P3RN的性能来验证EMAGAS的可行性。 实验结果表明，EMAGAS在信息熵，查询成本和匿名化时间方面具有不同程度的优势。

* 1. 贡献和影响

有两个主要贡献。首先，从商业角度，为提供LBS的移动商务建立了隐私保护服务框架，并且使用（K，L，P）匿名模型来描述移动用户的个性化隐私信息。第二，从技术角度出发，提出了一种名为EMAGAS的新的隐私保护算法，并通过实验验证了IMAGES的可行性和性能优势。

然而，实际上，存在许多LBS应用，并且不同的信息服务提供商通常具有不同的服务资源。单个信息服务提供商不能向移动用户提供全面和高质量的LBS建议。 所提出的MCA的服务框架不仅有助于保护移动用户的隐私，而且促进服务资源的共享。此外，鉴于智能移动电话的广泛使用和各种移动应用的普及，移动用户的隐私信息可以在任何位置和时间被公开。位置固定和移动电话号码绑定是各种移动商务应用中常用的特征，这可能导致移动用户的隐私信息的公开，诸如位置，身份和其他敏感信息。 隐私披露风险已成为影响移动服务进一步采用的制约因素之一。 因此，在理论和技术上探索用于保护移动用户的隐私的解决方案是重要的。

* 1. 限制和未来研究方向

我们的工作有一些限制。首先，必须设计用于MCA的所提出的隐私保护服务框架的应用架构，并且不仅要真实的道路网络而且还有真实的移动商务应用场景来进一步测试所提出的算法的性能。 第二，关于如何每个信息服务提供商基于匿名查询请求生成查询结果的列表的问题以及MCA平台如何细化接收到的结果来自多个信息服务提供商必须解决。

随着大数据时代的出现，移动商务为服务提供商提供了前所未有的实时精准营销的机会。 同时，也出现了与移动商务中的隐私问题相关的新挑战。 一方面，信息过载的问题和隐私信息披露的风险越来越严重（Feng et al。，2014）。 推荐系统是解决信息过载问题的有效解决方案。 在保护移动用户隐私信息的条件下提供高质量服务建议是一个具有挑战性的问题（Zhao et al。，2014）。

已有许多关于个性化推荐方法的研究。 众所周知，很少有关于移动商务的报告工作提出了基于隐私意识的服务推荐的整体解决方案。另一方面，基于位置大数据的敏感信息保护也是一个具有挑战性的问题。 位置大数据不仅涉及用户的位置隐私信息，而且意味着敏感信息，诸如个人习惯，健康状况和用户的社交位置。攻击者可以通过来自多个源的大数据的积累和关联来提取，集成，分析和发现关于移动用户的敏感信息。 必须深入研究防止攻击者通过位置数据和服务相关数据之间的连接获取移动用户的敏感信息的解决方案。

1. 杨璐, 田惠生, 贾明利,等. 一种高实时性的移动商务管理系统设计与实现[J]. 计算机工程, 2016, 42(1):292-299.
2. 李丹丹, 华蕊. 基于模糊综合评价法的移动商务身份认证评价[J]. 微电子学与计算机, 2016, 33(3):65-68.
3. Xue J, Liu X Y, Yang X C, et al. A Location Privacy Preserving Approach on Road Network[J]. Chinese Journal of Computers, 2011, 34(5):865-878.
4. 岑婷婷, 韩建民, 王基一,等. 隐私保护中K-匿名模型的综述[J]. 计算机工程与应用, 2008, 44(4):130-134.
5. Zhan F B, Noon C E. Shortest Path Algorithms: An Evaluation Using Real Road Networks[J]. Transportation Science, 1998, 32(1):65-73.
6. 李静, 韩建民. 一种含敏感关系社会网络隐私保护方法-(k,l)-匿名模型[J]. 小型微型计算机系统, 2013, 34(5):1003-1008.
7. 宋金玲, 赵威, 刘欣,等. k-匿名数据集的增量更新算法[J]. 计算机科学, 2010, 37(4):146-150.
8. Gangopadhyay A, Adya M. Protecting Sensitive Information in Electronic Commerce[M]// Doing Business on the Internet. 1999:77-86.
9. 陈爱东. K-匿名数据集的挖掘算法研究[D]. 东华大学, 2014.
10. Lu Q Z, Lin H L, Ge S T, et al. Wireless, remote-query, and high sensitivity Escherichia coli O157:H7 biosensor based on the recognition action of concanavalin A.[J]. Analytical Chemistry, 2009, 81(14):5846-50.
11. Hayes J, Troncoso C, Danezis G. TASP: Towards Anonymity Sets that Persist[C]// ACM. 2016.
12. 宋金玲, 刘国华, 黄立明,等. k-匿名方法中相关视图集和准标识符的求解算法[J]. 计算机研究与发展, 2009, 46(1):77-88.
13. 李晓燕, 朴春慧, 潘晓. DASS:路网上基于交换和合并的用户敏感信息保护方法[J]. 河北省科学院学报, 2014, 31(2):66-73.
14. 肖燕芳. 基于匿名区域变换的位置隐私保护模型与算法研究[D]. 华南理工大学, 2012.
15. 于娟. 数据发布中隐私保护的匿名模型及算法研究[D]. 浙江师范大学, 2010.
16. Guo K, Zhang Q. Fast clustering-based anonymization approaches with time constraints for data streams[J]. Knowledge-Based Systems, 2013, 46(1):95-108.
17. Baioco G B, Traina A J M, Traina C. An effective cost model for similarity queries in metric spaces[C]// ACM Symposium on Applied Computing. 2007:527-528.
18. Shou L, Shang X, Chen K, et al. Supporting Pattern-Preserving Anonymization for Time-Series Data[J]. IEEE Transactions on Knowledge & Data Engineering, 2013, 25(4):877-892.
19. Spinney J E. Mobile Positioning and LBS Applications[J]. Geography, 2003, 88(4):págs. 256-265.
20. Chan S, Sohn G. Indoor Localization Using Wi-Fi Based Fingerprinting and Trilateration Techiques for Lbs Applications[J]. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012, XXXVIII-4/C26:1-5.
21. Li D. Research on Applications of LBS Based on Electronic Compass Assisted GPS[C]// International Symposium on Information Engineering and Electronic Commerce. IEEE, 2009:599-602.
22. Zhang Y, Zhong S. A privacy-preserving algorithm for distributed training of neural network ensembles[J]. Neural Computing and Applications, 2013, 22(1):269-282.
23. 陈珂. 开放式环境下敏感数据安全的关键技术研究[D]. 浙江大学, 2007.
24. 张飞. 基于动态数据集的匿名化隐私保护技术研究[D]. 重庆交通大学, 2013.