面向位置推荐的差分隐私保护方法

夏 英 毛鸿睿 张 旭 裴海英
(重庆邮电大学计算机科学与技术学院 重庆 400065)

摘 要 位置推荐服务能使用户更容易地获得周边的兴趣点信息,但也会带来用户位置隐私泄露的风险。为了避免位置隐私泄露带来的不利影响,提出一种面向位置推荐服务的差分隐私保护方法。在保持用户位置轨迹与签到频率特征的前提下,基于路径前缀树及其平衡程度采用均匀分配和几何分配两种方式进行隐私预算分配,然后根据隐私预算分配结果添加满足差分隐私的Laplace噪音。实验结果表明该方法能有效保护用户位置隐私,同时通过合理的隐私预算分配能减少差分隐私噪音对推荐质量的影响。

关键词 位置推荐,差分隐私,隐私预算,Laplace噪音
中图法分类号 TP311 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.12.007

Differential Privacy Protection Method for Location Recommendation

XIA Ying MAO Hong-rui ZHANG Xu BAE Hae-young
(School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract Location recommendation service makes it easier for people to get surrounding information about Point of Interest (POI). However, there are some risks related to location privacy. In order to avoid the negative influence resulted from leaking location privacy, a privacy protection method for location recommendation service was proposed. On the premise of maintaining location trajectory and frequency characteristics of check-in, uniform distribution and geometry distribution were presented to control privacy budget allocation effectively based on path prefix tree (PP-Tree) and its balanced level, and thus the Laplace noise of differential privacy could be added according to the allocation result. Experiments indicate that this method can protect location privacy effectively. The impaction of differential privacy noise on the quality of location recommendation is reduced by reasonable privacy budget allocation.

Keywords Location recommendation, Differential privacy, Privacy budget, Laplace noise

1 引言

随着移动定位、移动互联网等技术的发展以及智能终端的普及,基于位置的服务(Location Based Service, LBS)越来越多地融入到人们的工作和生活中。通过LBS系统,用户能够不断提供或获得自己所处位置的相关信息,LBS系统也能够通过对用户历史位置数据的分析挖掘,为用户提供个性化推荐,如旅游路线、聚会地点等。

对大量位置数据进行分析挖掘是提供位置推荐服务的基础,但这可能引起用户位置泄露。当攻击者积累了一定的有关推荐系统的背景知识时,就可以推测出用户的位置信息,从而获取其兴趣爱好、生活模式等敏感信息[1]。例如,在K最近邻(K-Nearest Neighbor, KNN)攻击中,攻击者可以通过构造伪相似用户来推测用户位置[2]。位置推荐需要频繁且广泛的查询操作,传统的隐私保护方法在位置推荐服务中还存在不足,如基于加密的保护方法的复杂度较高、计算开销大[3-4];K-匿名及其扩展模型易受背景知识攻击[5-6]。差分隐私[7]是一种通过向查询或分析结果中添加适当噪音来达到隐私保护效果的方法,并且严格定义量化评估的方法。将差分隐私运用于推荐过程中能较好地保护用户位置隐私,比如文献[8]在推荐过程中通过对协同矩阵添加Laplace噪音来进行隐私保护;文献[9]提出的PNCF方法通过在评估成员间的相似度过程中添加Laplace噪音的方式来保护隐私,但噪音扰动一定程度上会影响最后的推荐质量。

为保护用户的位置隐私,减小噪音对位置推荐质量的影响,提出一种基于用户路径前缀树的差分隐私保护方法DPPT(Differential Privacy based on Path-Prefix-Tree)。DPPT结合用户位置轨迹和签到频率特征分配隐私预算,并根据分配结果添加Laplace噪音,同时利用差分隐私的序列性提出均匀分配和几何分配两种不同的隐私预算分配方式,

到稿日期:2016-10-11 返修日期:2016-11-12 本文受国家自然科学基金(41201378),重庆市自然科学基金(cstc2014kjrc-qnrc40002),重庆市教育科学技术研究项目(KJ1500431)资助。

夏 英(1972-),女,博士,教授,主要研究方向为数据库、数据挖掘、空间信息处理、云计算,E-mail:xiaying@cqupt.edu.cn;**毛鸿睿**(1992-),男,硕士生,主要研究方向为数据库、数据挖掘、位置隐私;**张 旭**(1981-),男,博士,副教授,主要研究方向为数据库、数据挖掘、云计算、大规模数据处理;**裴海英**(1948-),男,博士,教授,主要研究方向为数据库、空间信息处理、地理信息系统。

以有效合理地控制隐私预算分配。

2　相关定义

2.1　差分隐私的相关定义

差分隐私保护模型可以保证在某一数据集中添加或者删除一条记录对最终的结果几乎没有影响。其相关定义如下。

定义 1(ε-差分隐私[7])　给定数据集 D 和 $\bar{D}$,两者至多相差一条记录,即 $\| D\Delta\bar{D}\|\leqslant 1$;若查询函数 Q 在这两个数据集上的任意输出结果 O 满足下列不等式,则 Q 满足 ε-差分隐私。

$$Pr[Q(\bar{D})=O]\leqslant e^{\varepsilon}\times Pr[Q(D)=O]$$

其中,隐私预算参数 ε 表示隐私保护程度,ε 越小,隐私保护程度越高。差分隐私保护通过添加适当的噪音来实现。但是,过多的噪音会影响结果的可用性,相反,过少的噪音则无法提供足够的安全保障。而敏感度是决定噪音量的关键参数,它表示数据集中任一记录的变化对结果造成的最大改变。

定义 2(全局敏感度 ΔQ)　对于任意一个查询函数 $Q\rightarrow O$,其全局敏感度为:

$$\Delta Q=\max_{D,\bar{D}}\| Q(D)-Q(\bar{D})\|$$

在差分隐私中,ε 代表着隐私保护水平,ε 一旦耗尽,则会失去差分隐私保护的意义。因此,适当地运用差分隐私的序列性可以有效、合理地分配隐私预算。

性质 1(序列性)　设有函数 $Q_1,Q_2,\cdots,Q_n$,其隐私保护预算分别为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,则对于给定数据集 D,由这些函数构成的组合函数 $Q(Q_1(D),Q_2(D),\cdots,Q_n(D))$ 满足 ε-差分隐私,其中 $\varepsilon=\sum_{1}^{n}\varepsilon_i$。

在差分隐私中一般采取 Laplace 机制对数值型数据进行噪音处理,它是通过在结果中加入服从 Laplace 分布的随机噪音来实现 ε-差分隐私保护。

定义 3(Laplace 噪音机制)　给定数据集 D,设有函数 $Q\rightarrow O$,其全局敏感度为 ΔQ,则函数 $Q_{DP}=Q+Y$ 提供 ε-差分隐私,其中 $Y\sim Lap(\frac{\Delta Q}{\varepsilon})$,概率密度如下:

$$\rho(x)=\frac{1}{2b}\exp(-\frac{|x|}{b}),b=\frac{\Delta Q}{\varepsilon}$$

2.2　位置推荐的相关定义

目前,位置推荐在位置服务和推荐系统中有着广泛的研究,基于协同过滤的推荐方法也常被用于位置推荐[10-11]。协同过滤利用用户对项目的评分矩阵来计算相似度,通过比较相似度来选择出与其最相似的邻居集合。由于仅仅依靠用户的位置签到统计并不能准确评价用户对位置的兴趣度评分,因此结合用户签到的频率特征来计算用户对位置的兴趣度评分,以此构建用户-位置的评分矩阵并计算用户位置间的相似度。

定义 4　用户 u 的所有签到位置集合为 $S_L(u)=\langle l_1,l_2,\cdots,l_n\rangle$,$|S_L(u)|=n$。其中,该用户在位置 $l_i(1\leqslant i\leqslant n)$ 处的签到次数统计为 α_i。

定义 5　在位置 l 签到过的所有用户集合为 $S_U(l)=\langle u_1,u_2,\cdots,u_n\rangle$,$|S_U(l)|=m$。其中,用户 $u_i(1\leqslant i\leqslant m)$ 在位置 l 处的签到次数统计为 β_i。

定义 6　用户 u 对签到位置 l 的贡献度 $UFILF$(User Frequency-Inverse Location Frequency)和用户 u 受签到位置 l 的影响程度 $LFIUF$(Location Frequency-Inverse User Frequency)的形式如下,其中 N_L 表示位置总数,N_U 表示用户总数。

$$UFILF(u,l)=\frac{\alpha_l}{\sum_{i=1}^{|S_U(l)|}\beta_i}\times\log\frac{N_L}{|S_L(u)|}$$

$$LFIUF(u,l)=\frac{\alpha_l}{\sum_{i=1}^{|S_L(u)|}\alpha_i}\times\log\frac{N_L}{|S_U(l)|}$$

式(1)用于计算两个用户 u_i 和 u_j 的位置相似度 $sim(u_i,u_j)$。$S_L(u_i,u_j)$ 表示 u_i 和 u_j 共同签到的所有位置集合,$E_{u_i,l}$ 和 $E_{u_j,l}$ 表示用户 u_i 和 u_j 对位置 l 的兴趣度评分,$\bar{E}_{u_i,l}$ 和 $\bar{E}_{u_j,l}$ 表示用户 u_i 和 u_j 对所签到位置的兴趣度评分的平均值,其中 $E_{u,l}=UFILF(u,l)+LFIUF(u,l)$。

$$sim(u_i,u_j)=\frac{\sum_{l\in S_{L(u_i,u_j)}}(E_{u_i,l}-\bar{E}_{u_i,l})(E_{u_j,l}-\bar{E}_{u_j,l})}{\sqrt{\sum_{l\in S_{L(u_i)}}(E_{u_i,l}-\bar{E}_{u_i,l})^2}\sqrt{\sum_{l\in S_{L(u_j)}}(E_{u_j,l}-\bar{E}_{u_j,l})^2}}\quad(1)$$

3　面向位置推荐的差分隐私保护方法

位置推荐系统根据用户对位置的兴趣度评分来计算位置相似度,而兴趣度评分基于用户的签到位置集合,其中包含了轨迹序列以及签到频率特征。为了直观清晰地反映这种序列和特征,考虑采用路径前缀树(Path-Prefix-Tree,PP-Tree)来组织用户的签到位置信息。差分隐私的噪音添加主要根据隐私预算分配结果,随机分配隐私预算会扭曲用户原有的轨迹序列以及签到频率特征,影响位置相似度计算。因此,为减小噪音所带来的误差,在 PP-Tree 结构的基础上进行差分隐私预算分配,从而根据所分配的隐私预算对用户的位置签到统计添加满足差分隐私的 Laplace 噪音。

3.1　PP-Tree 数据结构

将用户的所有轨迹抽象为一个根节点为 root 的 PP-Tree。树中的每个节点描述一个位置,每个分支表示一条签到轨迹。用户的所有轨迹都以 root 为起点,将相同的子轨迹进行合并,同时对位置签到次数进行统计。图 1 给出了用户 u 的 PP-Tree 结构,其中左侧的列表表示该用户所有的轨迹序列。

序	路径
S_1	$l_1-l_2-l_3$
S_2	l_1-l_2
S_3	$l_3-l_2-l_1$
S_4	$l_1-l_2-l_4$
S_5	$l_3-l_2-l_5$
S_6	l_3-l_2
S_7	$l_1-l_2-l_4$
S_8	l_3-l_1

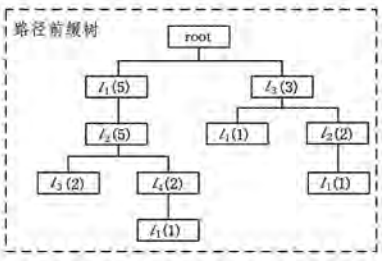

图 1　路径前缀树 PP-Tree 的结构

3.2　基于 PP-Tree 的隐私预算分配方案

相似用户签到同一个位置的概率是相近的,签到的轨迹也是相似的。为了保持用户间的相似度,隐私预算分配也遵循这种规律。考虑从根节点 root 出发,根据 PP-Tree 结构中每个位置的签到概率分布来分配隐私预算。在图 1 所示的 PP-Tree 中,用户 u 会以 0.625 的概率签到 l_1,以 0.375 的概

率签到 l_3。根据差分隐私的序列性，以 l_1 为根节点的轨迹集合分配隐私预算 $5\varepsilon/8$，以 l_3 为根节点的轨迹集合分配隐私预算 $3\varepsilon/8$。

PP-Tree 中抽象根节点 root 不是真实的签到位置，因此不会消耗隐私预算。但其他任何子结构的根节点都需要为其分配隐私预算。PP-Tree 的每层应受到相同程度的隐私保护，即每层所分配的隐私预算是相等的。但是当用户的轨迹序列长度差异较大时，PP-Tree 结构不平衡，子结构的高度差异较大，使得隐私预算分配倾斜。因此，本文提出均匀分配和几何分配两种方式进行隐私预算分配。当 PP-Tree 结构不平衡时，即子结构的高度差大于 1 时，采用几何分配；反之，采用均匀分配。

3.2.1 均匀分配

均匀分配根据 PP-Tree 结构使得每一层都分配相同的隐私预算。基于均匀分配的算法 UDPPT(Uniform Differential Privacy PP-Tree)见算法 1。由于 PP-Tree 中子树根节点具有唯一性，因此该节点的隐私预算为 $\bar{\varepsilon}=\frac{1}{h}\varepsilon$，其中 h 表示子树的高度，由最大轨迹长度决定。根据差分隐私序列性，剩下的隐私预算 $\frac{h-1}{h}\varepsilon$ 会按照签到下一位置的签到概率分布分配给当前根节点下的所有子结构。每个位置节点以〈*key*，*value*〉键值对的方式存于集合 *UMap* 中，*key* 表示位置节点标识，*value* 表示所分配的隐私预算。并且以队列 *UQueue* 的形式存储二元组〈v，ε_v〉，表示以节点 v 为根节点的 PP-Tree 可分配隐私预算为 ε_v。UDPPT 的计算复杂度与 *UQueue* 的大小呈正相关性，而 *UQueue* 的大小由用户 u 的签到位置总数 N_l 决定，因此 UDPPT 的时间复杂度为 $O(N_l)$。

算法 1 UDPPT

输入：隐私预算 ε，路径前缀树 PP-Tree，根节点 root
输出：位置隐私预算分配结果集合 UMap
1. 初始化 UMap 为空，UQueue 为空
2. 添加〈root，ε〉至 UQueue 中
3. While UQueue 不为空时 Do
4. 〈root，ε〉←UQueue 队列头部的元素出队列
5. height ← 以 root 为根节点的 PP-Tree 高度
6. If root∈UMap Then
7. ε_{root} ← UMap 中 root 节点的隐私预算
8. UMap∪〈root，ε_{root}+ε/height〉//替换原有键值对
9. Else
10. UMap∪〈root，ε/height〉//添加新的键值
11. EndIf
12. ε ← ε(height−1)/height
13. For 根节点 root 的每一个孩子节点 v
14. P_v ← 签到 v 的概率
15. 添加〈v，εP_v〉至 UQueue 中
16. EndFor
17. EndWhile

3.2.2 几何分配

几何分配使得每个子结构中的根节点所分配的隐私预算由该节点在当前结构中的签到统计频率决定。隐私预算与签到统计结果呈正相关性，即签到统计的频率越大，节点分配的隐私预算越多。基于几何分配的算法 GDPPT(Geometric Differential Privacy PP-Tree)见算法 2。每个位置节点以〈*key*，*value*〉键值对的方式存于集合 *GMap* 中。同样以队列的形式存储二元组〈v，ε_v〉，表示以节点 v 为根节点的 PP-Tree 可分配隐私预算为 ε_v。同样，GDPPT 的计算复杂度与 *GQueue* 的大小呈正相关性，时间复杂度也为 $O(N_l)$。

算法 2 GDPPT

输入：隐私预算 ε，路径前缀树 PP-Tree，根节点 root
输出：位置隐私预算分配结果集合 GMap
1. 初始化 GMap 为空，GQueue 为空
2. 添加〈root，ε〉至 GQueue 中
3. While GQueue 不为空时 Do
4. 〈root，ε〉 ← GQueue 队列头部的元素出队列
5. P_{root} ← root 节点的签到统计频率
6. If root∈GMap Then
7. ε_{root} ← GMap 中 root 节点的隐私预算
8. GMap∪〈root，ε_{root}+εP_{root}〉//替换原有键值对
9. Else
10. UMap∪〈root，εP_{root}〉//添加新的键值
11. EndIf
12. ε ← ε−εP_{root}
13. For 根节点 root 的每一个子节点 v
14. P_v ← 签到 v 的概率
15. 添加〈v，εP_v〉至 GQueue 中
16. EndFor
17. EndWhile

同一用户可能会在不同时间签到同一位置，因此同一位置会因多次签到而多次出现在 PP-Tree 结构中，并且由于签到位置轨迹不同，所处节点位置也会不同。因此，隐私预算分配可能会使同一位置分配到不同的隐私预算，需要按照位置标识将同一位置的不同分配结果进行合并。在图 2 所示的隐私预算分配结果中，假设 root 位于第 0 层，则第 1 层中位置 l_3 的隐私预算分配结果为 $\varepsilon/8$，第 3 层中的相同位置 l_3 的隐私预算分配结果为 $5\varepsilon/32$，因此该用户的签到位置 l_3 最后所分配的隐私预算为 $9\varepsilon/32$。差分隐私根据最后的隐私预算分配结果添加 Laplace 噪音，即 $\bar{c}(l_i)=c(l_i)+Laplace(\varepsilon_i)$。推荐系统用含有噪音的结果 $\bar{c}(l_i)$ 进行相似度计算，从而保护推荐过程中用户的位置隐私。

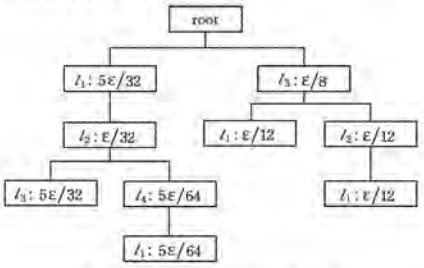

图 2 隐私预算分配结果示例

3.3 差分隐私保护下的位置推荐

位置推荐通过 *UFILF* 和 *LFIUF* 计算出用户 i 对位置 j 的兴趣度评分 E_{ij}，并构建用户-位置的兴趣度评分矩阵 M_E：

$$M_E=\begin{bmatrix} E_{11} & E_{12} & \cdots & E_{1j} \\ E_{21} & E_{22} & \cdots & E_{2j} \\ \cdots & \cdots & & \cdots \\ E_{i1} & E_{i2} & \cdots & E_{ij} \end{bmatrix}$$

再利用式(1)计算位置的相似度，构建相似性矩阵 M_i，其中 $Sim(i,j)$表示用户 u_i 和 u_j 的位置相似度。对每个用户而言，选择相似度最高的 k 个用户作为当前用户的相似邻居，并筛选出用户未签到但可能感兴趣的位置给予推荐。

$$M_i=\begin{bmatrix} Sim(1,1) & Sim(1,2) & \cdots & Sim(1,j) \\ Sim(2,1) & Sim(2,2) & \cdots & Sim(2,j) \\ \cdots & \cdots & & \cdots \\ Sim(i,1) & Sim(i,2) & \cdots & Sim(i,j) \end{bmatrix}$$

4　实验分析

4.1　实验方案

本实验选用真实的公共位置数据集 Gowalla，其中包含 107092 个用户和 1280970 个不同的位置地点，共 6442892 条用户位置签到记录[12]。为了直观地反映差分隐私对推荐质量的影响，实验中选取精确率(Precision)和召回率(Recall)作为评价推荐质量指标，其中有效推荐集合为推荐集合与测试集合的交集。

$$Precision=\frac{\text{有效推荐集合数量}}{\text{推荐集合数量}}$$

$$Recall=\frac{\text{有效推荐集合数量}}{\text{测试集合数量}}$$

为了客观分析 DPPT 方法在面向位置推荐服务时的可行性和效果，将该方法与基于差分隐私保护的 PNCF 方法[9]进行比较，严格控制差分隐私参数 ε，选取多个不同值进行对比。同时，为了验证两种隐私预算分配方式的有效性，分别采用基于均匀分配的 UDPPT 与基于几何分配的 GDPPT 进行实验。为了更直观地反映隐私保护对位置推荐的影响，在本文所提方法的基础上假设忽略差分隐私保护，将这种未进行隐私保护的位置推荐方法作为 Baseline，并与上述 3 种基于隐私保护的位置推荐方法进行对比。实验采用 5 折交叉验证，最后取精确率和召回率的平均值。所有实验方法均采用同一种相似性计算方法和 top-k(实验中 $k=10$)的方式筛选出相似度最高的候选位置集合。

4.2　实验结果分析

差分隐私预算对位置推荐的精确率和召回率的影响分别如图 3、图 4 所示。

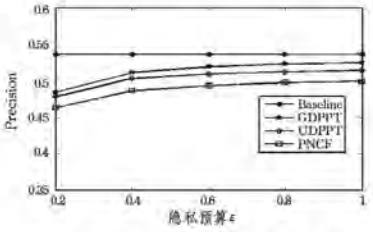

图 3　差分隐私预算对精确率的影响

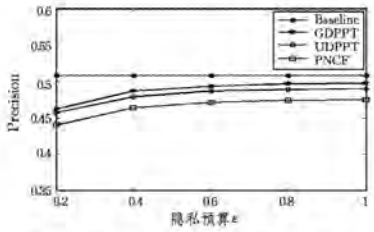

图 4　差分隐私预算对召回率的影响

实验结果表明，基于差分隐私保护的 UDPPT，GDPPT 和 PNCF 会造成一定程度的推荐质量损失，但是随着差分隐私参数 ε 的增大，其精确率和召回率逐渐回归于无隐私保护的水平。原因在于，隐私参数 ε 决定着差分隐私中噪音的大小，隐私参数 ε 越小表示隐私级别越高，噪音也越大。与 PNCF 相比，UDPPT 和 GDPPT 的精度损失较小，因为 UDPPT 和 GDPPT 基于 PP-Tree 结构添加 Laplace 噪音，保持了用户的原始轨迹以及签到频率特征，一定程度上维持了用户位置间的相似性，减小了因噪音添加造成的推荐质量损失；而基于相似度添加噪音的 PNCF 使用户位置间的相似性变得模糊。进一步发现，基于几何分配的 GDPPT 比基于均匀分配的 UDPPT 的精度损失更小，原因在于 UDPPT 分配隐私的预算受 PP-Tree 结构的影响。在本实验中，用户签到位置的轨迹长度差异较大，PP-Tree 结构不平衡，造成 UDPPT 中的隐私预算分配倾斜，而 GDPPT 独立于结构特征，根据位置签到频率来分配隐私预算并添加噪音。

结束语　本文为了在位置推荐服务中保护用户的位置隐私，提出了一种基于差分隐私的保护方法。该方法通过在用户的位置签到统计结果中添加适当的差分隐私噪音来达到隐私保护效果。考虑到差分隐私噪音会对位置推荐质量造成影响，利用 PP-Tree 来组织用户的签到位置信息，并基于 PP-Tree 及其平衡程度采用均匀分配和几何分配两种不同的隐私预算分配方式。通过分析真实位置数据集上的实验结果可知，该方法能有效保护用户位置隐私，并且合理选择隐私预算的分配方式能有效减小差分隐私噪音对推荐质量的影响。

参 考 文 献

[1] ZHANG X J, GUI X L, WU Z D. Privacy preservation for location-based services: A survey[J]. Journal of Software, 2015, 26(9): 2373-2395. (in Chinese)
张学军，桂小林，伍忠东. 位置服务隐私保护研究综述[J]. 软件学报，2015，26(9)：2373-2395.

[2] CALANDRINO J A, KILZER A, NARAYANAN A, et al. You Might Also Like: Privacy Risks of Collaborative Filtering[C]// IEEE Symposium on Security & Privacy. IEEE, 2011: 231-246.

[3] LIU S B, LI Y M, LIU M J. Privacy-preserving for Location-based Service over Encrypted Data Search[J]. Computer Science, 2015, 42(4): 101-105. (in Chinese)
刘树波，李艳敏，刘梦君. 基于密文检索的位置服务用户隐私保护方案[J]. 计算机科学，2015，42(4)：101-105.

[4] ZHAN J, HSIEH C L, WANG I C, et al. Privacy-Preserving Collaborative Recommender Systems[J]. IEEE Transactions on Systems Man & Cybernetics Part C, 2010, 40(4): 472-476.

[5] CORMODE G, PROCOPIUC C, SRIVASTAVA D, et al. Differentially Private Spatial Decompositions[C]// 2012 IEEE 28th International Conference on Data Engineering. IEEE Computer Society, 2012: 20-31.

[6] LIU B, FU Y, YAO Z, et al. Learning geographical preferences for point-of-interest recommendation[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2013: 1043-1051.

(下转第 57 页)

CIKM '08. Napa Valley, USA, 2008: 931-940.

[13] MASSA P, AVESANI P. Trust-aware recommender systems [C]// Proc. of RecSys '07. Minneapolis, MN, USA, 2007: 17-24.

[14] RESNICK P, IACOVOU N, SUCHAK M, et al. GroupLens: an open architecture for collaborative filtering of netnews [C]// Proceeding of the ACM Conference on Computer Supported Cooperative Work. 1994: 175-186

[15] SHI Y, LARSON M, HANJALIC A. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges[J]. ACM Computing Surveys (CSUR), 2014, 47(1): 1-45

[16] CACHEDA F, CARNEIRO V, FERNÁNDEZ D, et al. Comparison of collaborative filtering algorithms: limitations of current techniques and proposals for scalable, high-performance recommender system[J]. ACM Transactions Web, 2011, 5(1): 1-33.

[17] HUANG C G, YIN J, WANG J, et al. Uncertain neighbors' collaborative filtering recommendation algorithm[J]. Chinese Journal of Computers, 2010, 33(87): 1369-1377. (in Chinese)
黄创光，印鉴，汪静，等. 不确定近邻的协同过滤推荐算法[J]. 计算机学报，2010，33(87)：1369-1377.

[18] LUO X, OUYANG Y X, XIONG Z, et al. The effect of similarity support in K-nearest-neighborhood based collaborative filtering [J]. Chinese Journal of Computers, 2010, 33(8): 1437-1455. (in Chinese)
罗辛，欧阳元新，熊璋，等. 通过相似度支持度优化基于 K 近邻的协同过滤算法[J]. 计算机学报，2010，33(8)：1437-1455.

[19] XING C X, GAO F R, ZHAN S A, et al. A collaborative filtering recommendation algorithm in corporate with user interest change [J]. Journal of Computer Research and Development, 2007, 44(2): 296-301. (in Chinese)
邢春晓，高凤荣，战思南，等. 适应用户兴趣变化的协同过滤推荐算法[J]. 计算机研究与发展，2007，44(2)：296-301.

[20] STREHL A, GHOSH J, MOONEY R. Impact of similarity measures on web-page clustering[C]// Proceedings of the International Workshop on Artificial Intelligence for Web Search. 2000: 58-64.

[21] HERLOCKER J L, KONSTAN J A, BORCHERS A, et al. An algorithmic framework for performing collaborative filtering[C]// Proceedings of the Twenty Second Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 1999: 230-237.

[22] JAMALI M, ESTER M. Trustwalker: A random walk model for combining trust-based and item-based recommendation [C]// Proceedings of the fifteenth ACM SIGKDD international conference on knowledge discovery and data mining. ACM, 2009: 397-406

[23] SHARDANAND U, MAES P. Social information filtering: algorithms for automating word of mouth[C]// Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994: 210-217.

[24] POLATIDS N, GEORGIADIS C K. A multi-level collaborative filtering method that improves recommendations [J]. Expert Systems with Applications, 2016, 48: 100-110.

[25] LU J, SHAMBOUR Q, XU Y, et al. A web-based personalized business partner recommendation system using fuzzy semantic techniques[J]. Computational Intelligence, 2013, 29(1): 37-69.

[26] LUO H, NIU C, SHEN R, et al. A collaborative filtering framework based on both local user similarity and global user similarity[J]. Mach. Learn., 2008, 72(3): 231-245.

[27] LIU H, HU Z, MIAN A, et al. A new user similarity model to improve the accuracy of collaborative filtering[J]. Knowledge-Based Systems, 2014, 56(3): 156-166.

[28] WANG W, ZHANG G, LU J. Collaborative filtering with entropy-driven user similarity in recommender systems[J]. International Journal of Intelligent Systems, 2015, 30(8): 854-870.

[29] BHATTACHARYYA A. On a measure of divergence between two statistical populations defined by their probability distributions[J]. Bull. Calcutta Math. Soc., 1943, 35(1): 99-109.

[30] KAILATH T. The divergence and Bhattacharyya distance measures in signal selection[J]. IEEE Transactions Commun. Technol., 1967, 15(1): 52-60.

[31] NIELSEN F, BOLTZ S. The Burbea-Rao and Bhattacharyya centroids[J]. IEEE Transactions Inf. Theory, 2011, 57(8): 5455-5466.

[32] AHERNE F J, THACKER N A, ROCKETT P. The Bhattacharyya metric as an absolute similarity measure for frequency coded data[J]. Kybernetika, 1998, 34(4): 363-368.

[33] HUANG A. Similarity measures for text document clustering [C]// Proceedings of the Sixth New Zealand Computer Science Research Student Conference (NZCSRSC2008). Christchurch, New Zealand, 2008: 49-56.

(上接第 41 页)

[7] DWORK. Differential Privacy[C]// Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP). 2016: 1-12.

[8] MCSHERRY F, MIRONOV I. Differentially private recommender systems: building privacy into the net[C]// ACM Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM, 2009: 627-636.

[9] ZHU T, LI G, REN Y, et al. Differential privacy for neighborhood-based Collaborative Filtering [C]// IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE, 2013: 752-759.

[10] YE M, YIN P, LEE W C. Location recommendation for location-based social networks[C]// ACM Sigspatial International Symposium on Advances in Geographic Information Systems, ACM-GIS 2010. USA, 2010: 458-461.

[11] WANG H, TERROVITIS M, MAMOULIS N. Location Recommendation in Location-based Social Networks using User Check-in Data[C]// ACM Sigspatial International Conference on Advances in Geographic Information Systems. 2013: 374-383.

[12] CHO E, MYERS S A, LESKOVEC J. Friendship and mobility: user movement in location-based social networks [C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego, CA, USA, 2011: 1082-1090.