差分隐私在协同过滤算法中的应用研究

鲜征征[1,2] 李启良[2] 李 改[3] 李 磊[2]
(广东金融学院互联网金融与信息工程系 广州 510521)[1]
(中山大学数据科学与计算机学院 广州 510006)[2]
(顺德职业技术学院电子与信息工程学院 顺德 528333)[3]

摘 要 利用背景知识间接推导出个人隐私信息已成为Internet用户更担忧的问题,定义极为严格且可证明的差分隐私保护是目前解决该问题的最有效的隐私保护技术。Berlioz等将差分隐私保护技术应用于协同过滤算法之一的矩阵分解中,虽然提出了新的算法,但是缺少严格的证明过程。针对他们提出的算法,将补充相应的数学证明,然后将Chaudhuri等提出的目标函数加扰方法灵活应用于ALS目标函数中。此外,还给出一种差分隐私保护参数的选择方案。最后,在两个真实数据集上的实验验证结果表明,所提出的ALS目标函数加扰方法取得了更好的推荐效果。

关键词 协同过滤,个人隐私保护,差分隐私,矩阵分解
中图法分类号 TP311 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.05.015

Research on Application of Differential Privacy in Collaborative Filtering Algorithms

XIAN Zheng-zheng[1,2] LI Qi-liang[2] LI Gai[3] LI Lei[2]
(Department of Internet Finance and Information Engineering, Guangdong University of Finance, Guangzhou 510521, China)[1]
(School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China)[2]
(School of Electronics and Information Engineering, Shunde Polytechnic, Shunde 528333, China)[3]

Abstract Today, the problem of personal privacy inferred by attacker using some background knowledge has become the problems which the Internet users are more worried about. Differential privacy is defined very strictly and can be proved, and it is the most effective privacy protection technology to solve this problem at present. Berlioz et al[1] proposed to apply differential privacy into matrix factorization which is the one of the popular collaborative filtering methods. Several new algorisms were proposed by Berlioz et al, but they lacked the strict proof processes. In this paper, we firstly added the prove processes of these algorisms. And then the objective function with added noise method proposed by Chaudhuri was used into the objective function of ALS. In addition, a selection scheme of differential privacy was given. Finally, some experimental results on two real datasets show that our approach obtains better recommendation accuracy while protecting the personal privacy in the raw data.

Keywords Collaborative filtering, Personal privacy preserving, Differential privacy, Matrix factorization

推荐系统已成为Internet商户提供给用户在线服务的核心功能之一。推荐系统为何存在安全问题?例如在基于商品的协同过滤中,用户的交易会导致该商品与用户历史交易的其它商品的相似性增加。因此,攻击者可以跟踪一个与目标用户(攻击对象)相关的商品的相似性列表,然后确定其中的新商品。当一个类似商品出现在这些跟踪列表中,攻击者即可推断出被添加到目标用户记录的项目。由于众多隐患,更多的数据拥有者不愿提供个人数据,或从自己的数据中去除一些信息,甚至提供一些虚假信息,导致数据挖掘研究的数据源不够真实可靠,也就必然严重影响挖掘的结果。因此,如何能够确保个人隐私安全,消除用户疑虑,鼓励用户提供真实可靠的数据,已成为数据挖掘领域中一个亟待解决的问题。

关于隐私保护技术,Dwork等人于2006年提出了定义极为严格的、与背景知识无关的新型隐私保护模型——差分隐私保护(Differential Privacy, DP)[2],该模型可从根本上解决传统隐私保护模型的缺陷。

Berlioz等人于2015年研究了基于矩阵分解的协同过滤算法的隐私保护,首次提出将差分隐私保护技术应用于矩阵分解的潜在因子中[1]。但该研究工作存在一些不足之处,譬如对于提出的算法,未给出严格的数学证明、采用输出结果加扰方法未取得很好的推荐准确率等。本文将针对文献[1]的算法给出相应的证明且指正文中出现的问题,然后运用

到稿日期:2016-02-24 返修日期:2016-06-03 本文受广东省高校创新强校工程自主创新能力提升类培育项目,广东省自然科学基金(2016A030310018)资助。

鲜征征(1977—),博士生,讲师,主要研究方向为数据挖掘中的隐私保护、机器学习,E-mail:xianzhengzheng@126.com;李启良(1990—),男,硕士生,主要研究方向为数据挖掘、机器学习;李 改(1980—),男,博士,讲师,主要研究方向为推荐系统、数据挖掘;李 磊(1951—),男,教授,博士生导师,主要研究方向为数据库、数据挖掘、人工智能。

Chaudhuri 等人提出的目标函数加扰[3]的思想，对 ALS 矩阵分解优化算法进行目标函数加扰处理，接着给出一种差分隐私保护参数的选择方法，最后在 Movelens-1M 和 Netflix 数据集上进行实验验证，并进行相关比较。

1 相关工作

基于协同过滤的推荐系统的推荐准确率与用户的数量和用户对项目的评分密切相关[4]。

近年来，已有一些研究工作致力于加强推荐系统的隐私保护。Canny 提出分散存储用户的信息，并将用户划分为多个社区，那么攻击者就必须针对多个系统来进行攻击[5]；Polat 等提出通过随机加扰的方法为用户评分增加不确定性[6]，即攻击者攻击用户的数据只会导致部分数据被泄露；Nikolaedko 等提出将多方计算应用于矩阵分解，使得推荐学习仅仅使用项目信息，而不揭露用户的评分[7]。然而，这些技术都不能阻止攻击者根据矩阵分解后的结果来间接推导出用户的评分信息。Calandrino 等研究了由推荐带来的隐私风险问题，例如亚马逊、Hunch 等个性化推荐系统中存在的隐私泄露风险[8]。

由 Dwork 等人于 2006 年提出的差分隐私保护技术不考虑攻击者的背景知识和计算能力，并提供了可以证明的隐私保障[2]。近年来，关于差分隐私的研究工作主要分为两类：1)数据发布[9-10]；2)数据挖掘和机器学习[3,11]。将差分隐私应用于推荐系统方面典型的研究主要有如下几项工作。1)Machanavajjhala等研究了基于用户和商品之间的一个图链接的社区推荐系统的隐私保护问题。捕获到一个目标用户的每个商品的效用，即可获得一个效用向量，其目标是产生一个该商品的概率分布从而最大化用户的效用，同时又保护向量的隐私[12]。2)McSherry 等将差分隐私用于协同过滤推荐系统中[13]，对 item-to-item 协方差矩阵进行差分隐私处理，且将推荐系统分为学习阶段和预测阶段(即推荐)，阐述了在没有严重损失推荐的准确率的情况下实施差分隐私保护是可行的，但未考虑到潜在因子模型。3) Liu 等提出通过 SGLD (Stochastic Gradient Langevin Dynamics)将差分隐私用于贝叶斯后验采样，从而避免了高斯噪声对整个参数空间的影响[14]。4)Hua 等又提出首先阻止不受信任的推荐者使用用户的评分信息，然后采用目标函数加扰的方法实现差分隐私保护，并采用 SGD (Stochastic Gradient Descent)进行优化[15]。5)Berlioz 等将差分隐私保护技术应用于矩阵分解的潜在因子[1]，但是缺少严格的证明等问题。因此，本文针对 Berlioz 等的算法给出相关证明，并且将 Chaudhuri 等提出的目标函数加扰的思想[3]应用于矩阵分解中，与文献[14-15]不同，本文是对优化算法 ALS(Alternating Least Squares)的目标函数加扰，最后还提出一种差分隐私保护参数的选择方法。

2 理论基础

本节首先介绍协同过滤中矩阵分解模型及常用的两种优化算法，然后介绍近年来得到广泛研究的差分隐私保护模型及其实现机制。

2.1 协同过滤中的矩阵分解模型

协同过滤算法的数据核心是“用户-项目”评分矩阵，该矩阵通常是稀疏的，存在一些缺失的评分。矩阵分解 (Matrix Factorization，MF)的基本原理是将原始矩阵拆解为数个矩阵后再做内积，它在 Netflix 公司举办的推荐系统大赛上得到了广泛的应用，也是目前预测缺失评分最有效的方法之一。

2.1.1 目标函数

设 $R_{n\times m}$ 表示 n 个用户 m 个项目的评分矩阵，元素 r_{ui} 表示用户 u 对项目 i 的评分。经矩阵分解后，$R_{n\times m}$ 将被分解成用户因子隐含矩阵 $P_{n\times d}$ 和项目因子隐含矩阵 $Q_{d\times m}$，其中 d 通常取低维。再设 $\tilde{R}_{n\times m}$ ($\tilde{R}_{n\times m}=P_{n\times d}\times Q_{d\times m}$) 为 $R_{n\times m}$ 的近似矩阵。$\tilde{r}_{ui}$ ($\tilde{r}_{ui}=p_u^{\mathrm{T}}\cdot q_i$) 为 r_{ui} 的近似值，本文也称其为预测值。矩阵分解的目标函数为最小化正则化的平方误差，即式(1)：

$$\min_{P,Q}\sum_{r_{ui}\in R}e_{ui}^2+\lambda(\|p_u\|_2^2+\|q_i\|_2^2) \tag{1}$$

其中，$e_{ui}=r_{ui}-\tilde{r}_{ui}$，表示真实值与预测值之间的误差；$\|\cdot\|_2$ 表示取矩阵的二范数；λ 为正则化参数，用来防止过拟合。通常采用 SGD 和 ALS 两种优化方法来求解式(1)。

2.1.2 优化算法

(1)随机梯度下降法(SGD)

随机梯度下降是用梯度的无偏估计来代替梯度。针对矩阵分解的目标函数(见式(1))，采用 SGD 不断迭代地更新隐含矩阵 P 和 Q 中的每一个元素来求解，即式(2)和式(3)。

$$p_u\leftarrow p_u+\gamma(e_{ui}q_i-\lambda p_u) \tag{2}$$

$$q_i\leftarrow q_i+\gamma(e_{ui}p_u-\lambda q_i) \tag{3}$$

其中，p_u，q_i 分别为矩阵 P 和矩阵 Q 中的任一元素；λ 为式(1)中的正则化参数；γ 是一个常量，用来决定最小化误差梯度下降的速率，通常称为学习率。

(2)交替最小二乘法(ALS)

交替最小二乘法是通过不断迭代来求解的。其基本原理是：在每一迭代过程中，首先固定一个隐含因子矩阵(例如矩阵 P)，矩阵分解的目标函数(见式(1))就变成一个凸优化问题来求解另一因子矩阵 Q；然后固定另一个隐含因子矩阵(例如矩阵 Q)，同理可以求出矩阵 P。多次迭代直到收敛，即求得式(1)中最优的 P 和 Q。

2.2 差分隐私保护模型

关于隐私保护技术，k-anonymity[16]及其多个扩展模型虽然得到了广泛研究且在隐私保护领域影响深远，但均不能提供足够的安全保障。针对攻击者能从一个计算(例如计数查询)的输出结果推导出输入数据中的任何记录的问题，Dwork 等人于 2006 年提出了差分隐私保护模型。该模型与传统的隐私保护有着本质上的区别，定义了一个极为严格的攻击模型，并对隐私泄露风险给出了严谨、定量化的表示和证明。该方法虽基于数据失真技术，但所加入的噪声量与数据集大小无关，因此对于大型数据集，仅通过添加极少量的噪声即可达到高级别的隐私保护[17]。

2.2.1 差分隐私保护的定义

差分隐私保护模型的核心思想主要有两点：1)可以确保在输入数据集中插入或删除一条记录的操作不会影响任何计算(例如计数查询)的输出结果；2)该模型不关心攻击者所具有的背景知识，即使攻击者已经掌握除了一条记录之外的所有记录的敏感信息，该记录的敏感信息也无法被披露[18]。该模型的形式化定义如下。

定义 1[3,19](差分隐私) 给定两个至多相差一条记录的数据集 D_1 和 D_2(即 $|D_1\Delta D_2|\leqslant 1$)，如果对于一个设定的随机

算法 A（取值范围设为 $Range(A)$），在数据集 D_1 和 D_2 的任意输出结果 $O(O\in Range(A))$ 满足式(4)，则称算法 A 满足 ε-差分隐私。

$$\Pr[A(D_1)\in O]\leqslant e^{\varepsilon}\times\Pr[A(D_2)\in O] \tag{4}$$

其中，$\Pr[\cdot]$ 表示隐私被披露的概率，它由算法 A 的随机性所控制（A 的随机性与攻击者的背景知识无关）；ε 是隐私保护参数（也称隐私保护预算），用来表示隐私保护的力度，ε 越小意味着隐私保护力度越强。

2.2.2 实现差分隐私保护的方法

Dwork 在文献[3,19]中指出实现差分隐私保护的关键技术是添加噪声。如图 1 所示，f 表示某一计算函数（例如计数），$f(D_1)$ 和 $f(D_2)$ 分别表示关于仅有一条记录不同的两个数据集的计算结果，即可得到：算法 A 通过对输出结果的随机化来提供隐私保护，同时通过参数 ε 来保证在数据集中删除任一记录时算法输出同一结果的概率可以控制在 e^{ε} 之内，即不会发生显著变化。

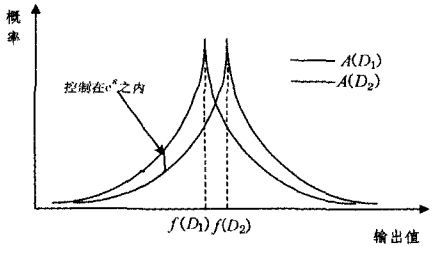


图 1　随机算法 A 在两个仅差一条记录数据集上的输出概率

通过添加噪声来满足差分隐私的算法与函数的敏感度和隐私保护参数 ε 直接关联。加入噪声过多会影响输出结果的可用性，噪声过少则导致无法提供足够的安全保障。

(1)函数的敏感度

函数的敏感度是决定加入噪声大小的关键参数。敏感度指函数从两个仅有一条记录不同的数据集中得到的输出结果的最大差别。函数全局敏感度(Global Sensitivity)是由函数自身决定的，与数据集无关。例如计数函数的全局敏感度为 1，只需加入少量噪声即可掩盖因一条记录被删而对函数输出结果所造成的影响。其形式化定义如下。

定义 2[3,19]（函数的全局敏感度）　对于任意一个函数 f：$D_1\rightarrow R^d$，d 表示函数 f 的维度，则函数 f 的 L_k-全局敏感度 $GS_k(f)$ 为：

$$GS_f=\max_{D_1,D_2}\|f(D_1)-f(D_2)\|_k \tag{5}$$

其中，$|D_1\Delta D_2|\leqslant 1$，$\|\cdot\|_k$ 表示 L_k-范数。

对于一些全局敏感度较大的函数（例如求平均值、中位数等），必须添加足够大的噪声才能保证隐私安全，这必将导致数据的可用性变差。因此，Nissim 等人提出并定义了局部敏感度[20]。局部敏感度是由函数自身和给定的数据集中的具体数据共同决定的。

(2)差分隐私的实现机制

将差分隐私保护技术应用于实践，实际就是设计并实现满足差分隐私保护要求的算法。针对不同的问题有不同的实现机制，拉普拉斯机制(Laplace Mechanism)和指数机制(Exponential Mechanism)[21]是两种最常见的差分隐私保护实现机制，前者适用于结果为数值型的保护，后者适用于结果为非数值型的保护。

1)拉普拉斯机制

拉普拉斯机制通过添加拉普拉斯随机噪声来实现差分隐私保护，其概率密度函数为：

$$f(x|\mu,b)=\frac{1}{2b}\exp(-|x-\mu|) \tag{6}$$

其中，μ 和 b 分别为变量 x 的位置参数和尺度参数，为方便获取噪声，设 $\mu=0$，记尺度参数 b 的 Laplace 分布为 $lap(b)$。

定义 3[25]（Laplace 机制）　给定一个数据集 D 和函数 f：$D\rightarrow R^d$，该函数的全局敏感度为 Δf，如果随机噪声 $Y\sim Lap(\Delta f/\varepsilon)$，则随机算法 $A(D)=f(D)+Y$ 满足 ε-差分隐私。

从定义 3 可知，加入的随机噪声与 Δf 成正比，与 ε 成反比。

2)指数机制

在许多实际应用中，某函数的结果可能是非数值型的(例如一种选择、一种方案)，为此，McSherry 等提出了针对非数值型结果的问题来实现差分隐私保护的指数机制。

定义 4[21]（指数机制）　给定一个数据集 D 作为随机算法 A 的输入，一个实体对象 $r(r\in Range(A))$ 作为 A 的输出，设 $q(D,r)$ 为可用性函数，Δq 为其全局敏感度。如果算法 A 以正比于 $\exp(\varepsilon q(D,r)/2\Delta q)$ 的概率从函数的输出域 $Range(A)$ 中选择并输出结果 r，则随机算法 A 满足 ε-差分隐私。

(3)差分隐私的特性

通常，一个复杂的隐私保护问题通常需要多次应用差分隐私保护技术，为了保证整个过程的隐私保护水平控制在给定的隐私保护预算之内，需用到差分隐私保护本身具有的两种重要的组合特性：序列组合特性和并行组合特性[22]。

性质 1（序列组合性）　设多个随机算法 $A_1,A_2,\cdots,A_n$，每个随机算法对应的隐私保护预算分别为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，且各自满足 ε_i-差分隐私，则这些算法构成的组合算法 A 在同一数据集 D 上满足 $\sum_{i=1}^{n}\varepsilon_i$-差分隐私。

性质 2（并行组合性）　设多个随机算法 $A_1,A_2,\cdots,A_n$，每个随机算法对应的隐私保护预算分别为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，且各自满足 ε_i-差分隐私，则这些算法构成的组合算法 A 对于不相交的数据集 $D_1,D_2,\cdots,D_n$ 满足 $\max\varepsilon_i$-差分隐私。

3　基于差分隐私的矩阵分解算法及证明

为了保护原始评分矩阵数据中的个人隐私(特别是能够间接推导的隐私问题)，Berlioz 等人[1]将差分隐私保护技术引入到矩阵分解中。矩阵分解过程大致可分为 4 个步骤：1)数据输入（原始评分矩阵）；2)具体分解处理（SGD 或 ALS）；3)输出分解后的用户矩阵和项目矩阵；4)评分预测（即做推荐）[1]。

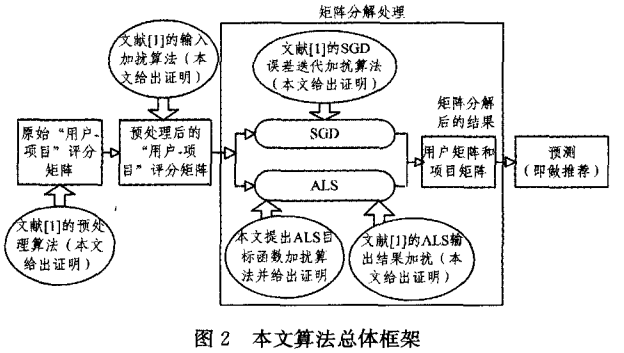


图 2　本文算法总体框架

文献[1]根据上述 4 个步骤，分别应用差分隐私保护技术，提出一系列差分隐私处理的新算法。若要满足差分隐私保护，则必须要有严格的证明，本文将对文献[1]的算法补充相关的数学证明，且针对个别算法中的错误加以说明和修正；此外，还提出了新的基于差分隐私保护的矩阵分解算法。图 2 示出了本文中几种算法的总体框架，详细描述请见 3.1—3.3 节。

3.1 进行差分隐私保护的预处理算法

文献[1]采用的预处理思想源于文献[18]，不同之处在于该文考虑了用户的平均评分，并将其并入到评分预测中。预处理涉及到的 3 个平均分为：全局平均分、项目平均分和用户平均分。

预处理过程主要包含 3 个步骤：1)计算带差分隐私保护的项目平均分；2)计算带差分隐私保护的用户平均评分；3)预测评分的恢复处理。

文献[1]对项目平均分做差分隐私保护处理的详细步骤参见原文算法 1 的描述。该算法首先对全局平均分加扰(总分加 Laplace 噪声后除以评分个数)，然后在此基础上进一步计算求得加扰后每个项目的平均分。因为用户评分属于 L_1-敏感度，因此评分的全局敏感度为：$\Delta r = r_{\max} - r_{\min}$。为了限制添加的噪声对只获得少量评分的项目带来的影响，在计算项目评分时设置了稳定参数 β_i。最后，为了保证加扰后的项目平均分有效，算法将其控制在$[r_{\min}, r_{\max}]$中。文献[1]中的算法 2 采用了类似算法 1 的方法来计算带差分隐私保护的用户平均分，但算法 2 中输入的评分矩阵变为 R'(R'中每个评分等于对应的原始评分减去算法 1 求得的项目平均分)，详细步骤参见原文。最后，文献[1]对预测评分进行恢复处理：最终预测评分＝项目平均分＋用户平均分＋矩阵分解后的预测评分。

针对文献[1]中的算法 1 和算法 2，本文给出定理 1 及其证明。

定理 1 给定一个评分矩阵 R、稳定参数 β_i 和隐私保护参数 ε，$\varepsilon=(\varepsilon_1+\varepsilon_2)$，$\varepsilon_1$ 用于全局平均分的隐私保护参数，ε_2 用于项目平均分的隐私保护参数，则算法 1 满足 ε-差分隐私。

证明：在算法 1 和算法 2 中实现差分隐私保护时是严格按照其实现机制来添加拉普拉斯噪声的，在此用到的是函数的全局敏感度。算法 1 中的第一步和第三步两次添加噪声的过程是串行的，再根据 2.1.2 节中的差分隐私保护序列组合特性，算法 1 满足$(\varepsilon_1+\varepsilon_2)$-差分隐私。同理，算法 2 满足$(\varepsilon_3+\varepsilon_4)$-差分隐私。

3.2 输入加扰的差分隐私保护算法

输入加扰的差分隐私保护的基本原理：对每个经过算法 1 和算法 2 预处理后的评分添加相应的拉普拉斯噪声，然后进行矩阵分解求得隐含因子矩阵。具体过程见文献[1]中算法 3 的描述。本文给出算法 3 的定理及证明。

定理 2 给定一个经过算法 1 和算法 2 预处理后的评分矩阵 R，隐私保护参数 ε，$\varepsilon=\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+\varepsilon_5$，其中，$\varepsilon_1$ 和 ε_2 是用于预处理算法 1 中的隐私保护参数，ε_3 和 ε_4 是用于预处理算法 2 中的隐私保护参数，ε_5 则用于算法 3，算法 3 满足 ε-差分隐私。

证明：因为在算法 3 的第一步中输入数据严格按照差分隐私实现机制添加拉普拉斯噪声，所以算法 3 满足 ε_5-差分隐私。由于这一步中被添加噪声的数据经过算法 1、算法 2 的预处理，因此算法 3 实际上满足 $\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+\varepsilon_5$-差分隐私。为了统一和简化隐私保护参数，文献[1]从整体上给出一个隐私保护预算参数 ε，然后将这个总的 ε 合理地分割为 5 份(具体分割请见实验结果部分)，再由序列组合特性可得算法 3 提供 ε-差分隐私。

3.3 矩阵分解过程中的差分隐私保护算法

通常采用 SGD 或 ALS 求解矩阵分解模型的目标函数(见 2.1.1 节式(1))。文献[1]分别针对 SGD(算法 4)和 ALS(算法 5)优化算法提出了相应的差分隐私保护算法。本文将给出这两个重要算法的详细证明过程，并对文献[1]中出现的错误加以说明和更正。

3.3.1 SGD 差分隐私保护

在 SGD 中引入差分隐私保护的主要思想是对每一次迭代中的误差添加拉普拉斯噪声，也称为梯度加扰。考虑到差分隐私的序列组合特性(见 2.1.2 节)，在每一次迭代中分配的隐私保护预算应该为 ε_5/k(k 为迭代次数)。具体描述见文献[1]中的算法 4。下面给出算法 4 的定理和证明。

定理 3 给定一个经过算法 1 和算法 2 预处理后的评分矩阵 R，以及用于 SGD 的多个参数，包括：矩阵分解隐含特征矩阵的特征个数 d，梯度下降的学习率 γ，正则化参数 λ，梯度下降迭代次数 k，评分误差的上界 $e_{\max}$、下界 $e_{\min}$和隐私保护参数 ε，$\varepsilon=\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+\varepsilon_5$，其中，$\varepsilon_1$ 和 ε_2 是用于预处理算法 1 中的隐私保护参数，ε_3 和 ε_4 是用于预处理算法 2 中的隐私保护参数，ε_5 用于 SGD 梯度加扰，算法 4 满足 ε-差分隐私。

证明：首先，算法 4 的输入是经过算法 1 和算法 2 预处理后的数据集，并且这两个算法都是满足差分隐私的(定理 1)。然后，算法 4 中每一次迭代过程中分配的隐私保护预算为 ε_5/k，根据差分隐私的序列组合特性，经过 k 次迭代后，算法 4 的隐私保护预算为：$\varepsilon=\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+(\varepsilon_5/k)*k$，因此，算法 4 满足 ε-差分隐私。

3.3.2 ALS 差分隐私保护

根据 ALS 求解矩阵分解目标函数(式(1))的原理(见 2.1.2节)，式(1)将变成凸优化问题，即式(7)和式(8)。

$$J_Q(p_u,R)=\sum_{R_u}(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2+n_u\lambda\parallel p_u\parallel_2^2 \tag{7}$$

$$J_P(q_i,R)=\sum_{R_i}(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2+n_i\lambda\parallel q_i\parallel_2^2 \tag{8}$$

采用 ERM(Empirical Risk Minimization)的思想求解式(7)和式(8)，即变成求解式(9)和式(10)。

$$p_u(R,Q)=\underset{p_u}{\operatorname{argmin}}\ J_Q(p_u,R) \tag{9}$$

$$q_i(R,P)=\underset{q_i}{\operatorname{argmin}}\ J_P(q_i,R) \tag{10}$$

(1)目标函数输出结果加扰

Chaudhuri 等人[3]提出了目标函数加扰和函数输出结果加扰的差分隐私处理方案，且将这两种方法应用于逻辑回归和支持向量机，文献[1]中的算法 5 通过对变形后的 ALS 目标函数(9)和(10)的输出结果加扰来实现差分隐私保护。本文首先对该算法中的错误加以说明，然后给出该算法的证明过程。

该算法中，当固定矩阵 Q 再求解矩阵 P 时，为式(9)的输出结果添加 Laplace 噪声，生成的随机噪声向量 b 的概率密度为：

$$f(b)\propto\exp\left(-\frac{\varepsilon\parallel b\parallel_2}{2k}\cdot\frac{n_u\lambda}{p_{\max}\Delta r}\right)$$

由于式(9)的敏感度是 L_2-敏感度，且全局敏感度为：

$$\Delta p_u=\frac{2q_{\max}\Delta r}{n_u\lambda}$$

其中，$q_{\max}$表示 $\| q_i \|_2$ 的上界，Δr 表示最高评分与最低评分的差，因此根据 Laplace 噪声机制，生成的随机噪声向量 b 的概率密度为：

$$f(b)\propto\exp(-\frac{\varepsilon\| b\|_2}{2k}\cdot\frac{n_u\lambda}{2q_{\max}\Delta r})$$

同理，$p_{\max}$表示 $\| p_u \|_2$ 的上界，式(10)的全局敏感度为：

$$\Delta q_i=\frac{2p_{\max}\Delta r}{n_i\lambda}$$

那么生成的随机噪声向量 b 的概率密度为：

$$f(b)\propto\exp(-\frac{\varepsilon\| b\|_2}{2k}\cdot\frac{n_i\lambda}{2p_{\max}\Delta r})$$

更改后的算法 5 的详细描述如下。

算法 1　DPALSOutput

输入：算法 1 和算法 2 做了预处理的用户评分矩阵 $R_{n\times m}=\{r_{ui}\}$；矩阵分解隐含特征矩阵的特征个数 d；正则化参数 λ；ALS 算法迭代次数 k；$\| q_i \|_2$ 上界 $q_{\max}$；$\| p_u \|_2$ 上界 $p_{\max}$；ALS 输出加扰的差分隐私保护预算参数 ε_5

输出：潜在因子用户矩阵 $P_{n\times d}$和项目矩阵 $Q_{d\times m}$

1. 初始化随机潜在因子矩阵 P 和 Q
2. for k 次迭代 do
3.　for 固定矩阵 Q，对每一个用户 u do
4.　　产生随机噪声向量 b，且概率密度为：

$$f(b)\propto\exp(-\frac{\varepsilon\| b\|_2}{2k}\cdot\frac{n_u\lambda}{2q_{\max}\Delta r})$$

5.　　$p_u\leftarrow\underset{p_u}{\operatorname{argmin}} J_Q(p_u,R)+b$

6.　　if $\| q_i \|_2>q_{\max}$ then $p_u\leftarrow p_u\cdot\frac{q_{\max}}{\| q_i \|_2}$

7.　end for
8.　for 固定矩阵 P，对每一个项目 i do
9.　　产生随机噪声向量 b，且概率密度为：

$$f(b)\propto\exp(-\frac{\varepsilon\| b\|_2}{2k}\cdot\frac{n_i\lambda}{2p_{\max}\Delta r})$$

10.　　$q_i\leftarrow\underset{q_i}{\operatorname{argmin}} J_P(q_i,R)+b$

11.　　if $\| p_u \|_2>p_{\max}$ then $q_i\leftarrow q_i\cdot\frac{p_{\max}}{\| p_u \|_2}$

12.　end for
13. end for
14. 返回潜在因子矩阵 $P_{n\times d}$和 $Q_{d\times m}$

下文将补充上述算法满足差分隐私保护的证明过程，详细证明过程见推论 1 和定理 4。该算法中，添加的随机噪声由推论 1 来计算。

推论 1　如果 $N(\cdot)=\| P_u \|_2$ 是 1-强凸且可导的，损失函数 $\ell=(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2$ 是凸函数且可导的，那么 $J_Q(P_u,R)$ 的 L_2-敏感度至多为 $2q_{\max}\Delta r/n_u\lambda$。

证明：首先，设两个仅有一个评分不同(在此设最后一个评分不同)的评分矩阵为：

$$R=\begin{pmatrix}r_{11}&\cdots&r_{1m}\\ \vdots&\ddots&\vdots\\ r_{n1}&\cdots&r_{nm}\end{pmatrix},R'=\begin{pmatrix}r_{11}&\cdots&r_{1m}\\ \vdots&\ddots&\vdots\\ r_{n1}&\cdots&r'_{nm}\end{pmatrix}$$

令：$G(p_u)=J_Q(p_u,R)$，$p_{u1}=\underset{p_u}{\operatorname{argmin}} J_Q(p_u,R)$，$p_{u2}=\underset{p_u}{\operatorname{argmin}} J_Q(p_u,R')$ 和 $g(p_u)=J_Q(p_u,R')-J_Q(p_u,R)=(r'_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2-(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2$。再令：$g(p_u)=(r'_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2-(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2$。

因为 $N(\cdot)=\| P_u \|_2$ 是 1-强凸，损失函数 $\ell=(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2$ 也是凸函数，所以 $G(P_u)=J_Q(P_u,R)$则是 $n_u\lambda$-强凸函数。

又因为 $N(\cdot)=\| P_u \|_2$ 和损失函数 ℓ 都是可导的，所以 $G(P_u)$和 $g(P_u)$也是可导的。接着，对 $g(P_u)$求梯度：

$$\begin{aligned}\nabla g(p_u)&=-2(r'_{ui}-p_u^{\mathrm{T}}\cdot q_i)q_i+2(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)q_i\\&=2q_i(r_{ui}-r'_{ui})=2q_i\Delta r\end{aligned}$$

可以得到：$\| \nabla g(p_u) \|=2\Delta r\| q_i^{\mathrm{T}} \|\leqslant 2q_{\max}\Delta r$。

因此，$J_Q(Pu,R)$的 L_2-敏感度小于等于 $2q_{\max}\Delta r/n_u\lambda$，推论得证。

定理 4　如果 $N(\cdot)=\| P_u \|_2$ 是 1-强凸且可导的，损失函数 $\ell=(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2$ 是凸函数且可导，那么当 $|\ell(\cdot)|\leqslant 1$ 时，算法 5 满足 ε-差分隐私，并且 $\varepsilon=\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+\varepsilon_5$。

证明：由推论 1 可知，如果关于 $N(\cdot)=\| P_u \|_2$ 和损失函数 $\ell=(r_{ui}-p_u^{\mathrm{T}}\cdot q_i)^2$ 的条件满足，那么 $J_Q(Pu,R)$的 L_2-敏感度至多为 $2q_{\max}\Delta r/n_u\lambda$。

又因为随机噪声向量 b 的概率密度函数为 $v(b)=\frac{1}{\alpha}e^{-\beta\| b\|}$，且 $\beta=\frac{n_u\lambda\varepsilon_5}{2q_{\max}\Delta r}$。再设 $R_{n\times m}$和 $R'_{n\times m}$为两个仅相差一个评分的邻矩阵，则有：

$$\frac{g(p_u|R)}{g(p_u|R')}=\frac{v(b_1)}{v(b_2)}=e^{-\frac{n_u\lambda\varepsilon_5}{2q_{\max}\Delta r}\| b_1-b_2\|}$$

(b_1，b_2 分别是 $g(p_u|R)$和 $g(p_u|R')$的随机噪声)。

如果 p_{u1}和 p_{u2}是未做隐私保护的函数 $J_Q(\cdot)$的两个解，那么当输入 $R_{n\times m}$和 $R'_{n\times m}$数据集时，有 $b_1-b_2=p_{u1}-p_{u2}$。再从推论 1 和三角不等式可得：

$$\| b_1 \|-\| b_2 \|\leqslant\| b_1-b_2 \|\leqslant\| p_{u1}-p_{u2} \|\leqslant\frac{2q_{\max}\Delta r}{n_u\lambda}$$

考虑到 b_1，b_2 的对称性，得到$\frac{v(b_1)}{v(b_2)}\leqslant e^{\varepsilon_5}$，即满足差分隐私定义。

隐私保护预算 $\varepsilon=\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4+\varepsilon_5$，其中$(\varepsilon_1+\varepsilon_2)$用于算法 1，$(\varepsilon_3+\varepsilon_4)$用于算法 2，$\varepsilon_5$ 用于 ALS，隐私保护预算的分配见第 4 节。

最后，因为差分隐私保护具有的序列组合性质，所以算法 5 是满足 ε-差分隐私的。同理，可以固定潜在因子矩阵 P 来优化求解 Q，算法 6 同样满足 ε-差分隐私。定理 4 得证。

(2)目标函数加扰算法(DPALSObjective)

目标函数加扰是 Chaudhuri 等人[3]提出的新方法，该方法理论上被证明是满足差分隐私保护的，实验结果也表明其可以较好地权衡隐私保护和推荐效果。由于 ALS 的目标函数(式(9)、式(10))满足文献[3]的算法 2，因此本文得到新的 ALS 目标函数加扰算法 DPALSObject，该算法应用了文献[3]的算法 2，并未引入任何新的步骤，本文只是针对 ALS 求出了相应的参数，考虑到篇幅问题，在此不重复给出算法步骤(具体请参见文献[3]的算法 2)，仅阐述具体应用于 ALS 的方法及步骤。

第一步　已知式(7)是求解隐含因子矩阵 P 的目标函数。

第二步　算法 DPALSObject 的主要思想是给 ALS 的目标函数(式(7))加扰，变为式(11)，其中 b 表示噪声向量：

$$J_Q^{priv}(p_u,R)=J_Q(p_u,R)+\frac{1}{n}b^{\mathrm{T}}p_u\tag{11}$$

第三步　最小化式(11)，可得：

$$p_u^{priv}=\underset{p_u}{\operatorname{argmin}} J_Q^{priv}(p_u,R)+\frac{1}{2}\Delta\|p_u\|_2 \tag{12}$$

第四步 针对式(12)分别对 P_u 求偏导。

当 $\forall 1\leqslant u\leqslant m, 1\leqslant k\leqslant d$ 时,即有:

$$\frac{1}{2}\frac{\partial p_u^{priv}}{\partial p_{uk}}=\sum_i(p_u^{\mathrm{T}}\cdot q_i-R_{ui})q_{ik}+\lambda n_u p_{uk}+\frac{1}{n}b_k+\frac{1}{2}\Delta p_{uk}$$

I 是 $d\times d$ 的单位矩阵,又因为:

$$\frac{1}{2}\frac{\partial p_u^{priv}}{\partial p_{uk}}=\frac{1}{2}\left(\frac{\partial p_u^{priv}}{\partial p_{u1}},\cdots,\frac{\partial p_u^{priv}}{\partial p_{ud}}\right)$$

$$=p_u[Q^{\mathrm{T}}Q+(\lambda n_u+\frac{1}{2}\Delta)I]-R_uQ+\frac{1}{n}b$$

第五步 分别求解矩阵 P 和矩阵 Q。首先固定矩阵 Q 并令 $\frac{\partial p_u^{priv}}{\partial p_{uk}}=0$,得到:

$$p_u=(R_uQ-\frac{1}{n}b)[Q^{\mathrm{T}}Q+(\lambda n_u+\frac{1}{2}\Delta)I]^{-1}$$

其中,$n_u=|R_u|$,$R_u=\{r_{vi}\in R|v=u\}$,m 表示评分矩阵中的用户数。

类似地,固定矩阵 P,可以求得 Q:

$$q_i=(R_iP-\frac{1}{n}b)[P^{\mathrm{T}}P+(\lambda n_i+\frac{1}{2}\Delta)I]^{-1},\forall 1\leqslant i\leqslant n$$

其中,$n_i=|R_i|$,$R_i=\{r_{vi}\in R|v=i\}$,n 表示评分矩阵中的项目数。

第六步 设算法 DPALSObject 中用于 ALS 的差分隐私保护预算为 ϵ_5,根据文献[3]中的算法 2 推导出应用于 ALS 时的参数 C 的取值为 2(具体推导过程与文献[3]中在逻辑回归和支持向量机的应用推导雷同,不再赘述)。

第七步 算法 DPALSObject 满足 ϵ-差分隐私(因为文献[3]的算法 2 已经被证明满足 ϵ-差分隐私)。

4 实验结果及分析

针对文献[1]提出的 5 个算法,Berlioz 等人选用了 Movielens中不同大小的数据集作为实验数据,考虑到算法对数据集的拟合问题,本文选用 Movielens-1M 和截取部分 Netflix 数据集进行实验验证。在实验评估指标上,与文献[1]相比,本文增加了 MAE 评估指标。需要说明的是,实验中提到的算法 1—算法 4 是指文献[1]中的 4 个算法,算法 5 是修正后的算法。

4.1 实验数据和评估指标

4.1.1 实验数据

本文实验数据选用两个数据集:Movielens-1M 和从 Netflix 数据集中随机截选的部分数据。它们的统计属性分别如表 1 所列。

表 1 两个数据集的统计属性

属性名	Movielens-1M 统计值	截取的 Netflix 统计值
用户数	6040	4996
电影数	3952	3999
密度/%	4.19	0.19
平均评分	3.5816	3.5956
评分的方差	1.2479	1.2208
每个用户参与评分的平均电影数	165.6	7.6
每部电影参评的平均用户数	253	9.5

4.1.2 实验评估指标

本文实验采用 10-折交叉验证来训练和预测(训练集:验证集=9:1)。在评估推荐准确率的过程中采用了根均方误差(RMSE)和平均绝对值误差(MAE)两个指标,RMSE 和 MAE 越小意味着预测结果越准确。RMSE 和 MAE 的计算方法见式(13),其中,$|R|$ 表示评分的数目。需注意的是,这里的评分指有效的评分,不包括那些缺失的评分。

$$RMSE=\sqrt{\frac{\sum_R(r_{ui}-\tilde{r}_{ui})^2}{|R|}},MAE=\frac{\sum_R|r_{ui}-\tilde{r}_{ui}|}{|R|} \tag{13}$$

4.2 数据的预处理

本文实验采用文献[1]中的算法 1 和算法 2 对原始评分数据集进行预处理。参考文献[1],各算法中差分隐私保护预算参数 ϵ 的分配方案如表 2 所列,总体来说 0.3ϵ 用于数据的预处理,0.7ϵ 用于矩阵分解。

表 2 差分隐私保护预算参数 ϵ 的分配方案

隐私保护参数的作用	分配比例
计算全局评分隐私保护参数	0.02ϵ
计算项目平均分隐私保护参数	0.14ϵ
计算用户平均分隐私保护参数	0.14ϵ
输入加扰、SGD 误差加扰、ALS 目标函数加扰和输出结果加扰算法用到的差分隐私保护参数	0.7ϵ

4.3 实验结果及分析

4.3.1 实验参数设置

实验中,算法的各个参数的具体选择如下。

1)稳定参数的设置遵循文献[1]中的算法 1 和算法 2,取值分别为:$\beta_i=15$,$\beta_u=20$;

2)依据经验,隐含特征矩阵的特征个数通常取低维,取值为:$d=5$;

3)依据经验,正则化参数取值为:$\lambda=0.125$;

4)SGD 和 ALS 求解迭代次数的设置方法:算法事先给定一个上限($k=20$),当误差变化小于 0.0001 时迭代停止;

5)依据经验,SGD 的学习率取值为:$\gamma=0.001$;

6)算法 DPALSObject 中参数 C 的取值是根据文献[3]中算法 2 提供的方法来求取的,$C=2$。

4.3.2 实验结果及分析

本小节首先对实验结果中图标的含义做简单介绍,然后对图 3—图 5 进行详细的分析。

1)* Baseline:不做任何差分隐私处理的矩阵分解,* 代表 SGD 或 ALS。

2)p*:表示按文献[1]中算法 1 和算法 2 做预处理。

3)pInput*:表示输入加扰,* 代表 SGD 或 ALS。

4)pDPSGD:表示 SGD 误差加扰。

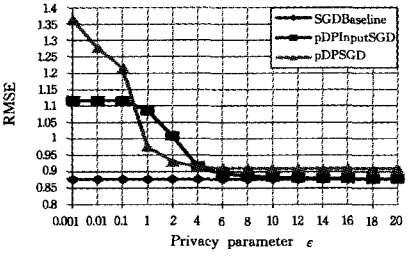
5)pALSDPOuput:表示 ALS 输出结果加扰。

6)pALSDPObject:本文提出的算法,表示 ALS 目标函数加扰。

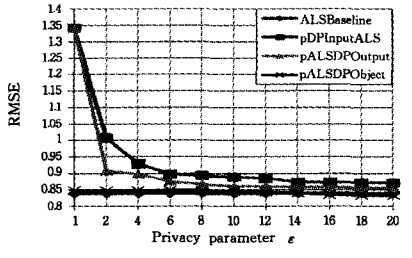
图 3 示出了分别对 SGD 和 ALS 算法进行差分隐私保护后与各自的 Baseline 进行比较的结果。限于篇幅,在此仅给出 RMSE 的结果,MAE 的结果与图 3 的趋势相似。图 3(a)和图 3(b)是在 Moivelens-1M 数据集上的结果,图 3(c)和图 3(d)是在截取的 Netflix 数据集上的结果。

从图 3 可以看出,无论在哪个数据集上,SGD 和 ALS 的不同的差分隐私保护算法在一定范围内的结果是可接受的。对于 SGD 而言,当 $\epsilon>1$ 时,预测的结果更加接近它的 Baseline。而对于 ALS 而言,本文采用的 ALS 目标函数加扰方法的结果最接近它的 Baseline。

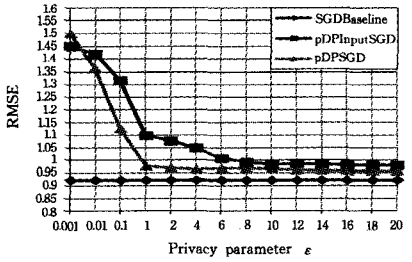
图 4 示出了 2 种 SGD 和 3 种 ALS 的差分隐私保护算法在不同数据集上的比较结果。



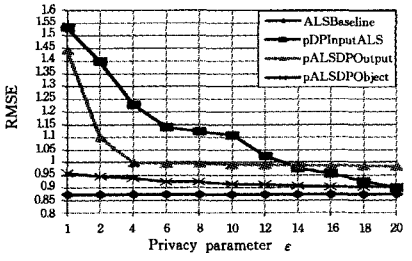
(a) Moivelens-1M 数据集上 2 种 SGD 的差分隐私保护方法比较



(b) Moivelens-1M 数据集上 3 种 ALS 的差分隐私保护方法比较



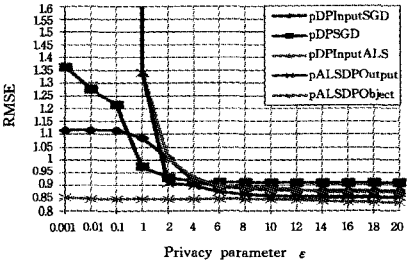
(c) 截取的 Netflix 数据集上 2 种 SGD 的差分隐私保护方法比较



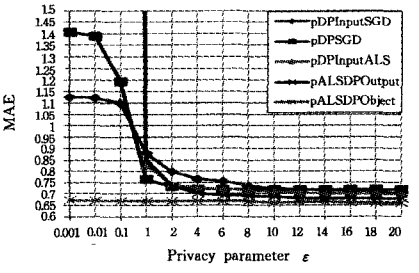
(d) 截取的 Netflix 数据集上 3 种 ALS 的差分隐私保护方法比较

图 3　对 SGD 和 ALS 算法进行差分隐私保护后与各自的 Baseline 进行比较的结果

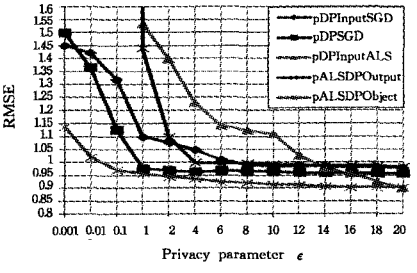
从图 4 可以看出，本文提出的 ALS 目标函数加扰的算法效果最佳，且随着 ε 变化也最稳定，这是因为该方法没有对原始数据加入过多的噪声。由于 ALS 方法自身优于 SGD，因此做了相关差分隐私保护处理后最终的结果还是凸显优势，当 $\varepsilon<2$ 时，ALS 输入加扰和输出结果加扰的方法几乎不能用于推荐，前者是因为对原始输入的数据加噪太多从而影响训练的模型，后者是因为对分解后的矩阵进行加扰时，ε 越小而噪声越大，那么两个加了较大噪声的矩阵再做内积得到的结果必定大大偏离真实值。由于 SGD 的每次迭代更新与误差相关，而 ALS 的每次迭代直接与训练的数据集有关，因此 ALS 对数据的变化表现得更为明显。各种算法在 Movielens-1M 数据集上的效果略优于截取的 Netflix，因为截取的数据集的训练样本比 Movielens-1M 少，且更加稀疏，所以预测的准确率跟数据集大小和稀疏度有着密切的关系。上述实验结果表明，本文的 ALS 目标函数加扰方法在隐私和推荐两个方面略优于文献[1]的算法。由于差分隐私保护性能问题至今仍是该领域研究的难点之一，通常认为在满足差分隐私保护定义的前提下，隐私保护参数取值越小则保护力度越大。换句话说，在推荐准确率相当的情况下，隐私保护参数取值越小，保护性能则越好。但从实际出发，隐私保护参数可以获得一个相对有效的范围。本文针对差分隐私保护参数给出一种实用的选择方案。



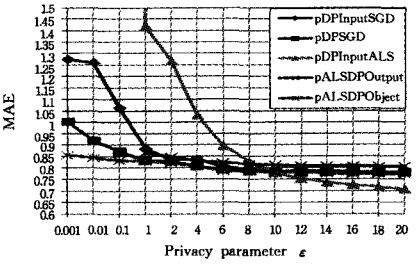
(a) 多种算法在 Movielens-1M 数据集上的 RMSE 比较结果



(b) 多种算法在 Movielens-1M 数据集上的 MAE 比较结果



(c) 多种算法在截取的 Netflix 数据集上的 RMSE 比较结果



(d) 多种算法在截取的 Netflix 数据集上的 MAE 比较结果

图 4　几种差分隐私保护算法在不同数据集上的不同评价指标的比较结果

4.3.3　差分隐私保护参数的选择

本文给出一种差分隐私保护参数 ε 的选择方案，基本原理及处理步骤简述如下：

第一步　确定被推荐的用户对象；

第二步　计算未作任何差分隐私处理时给第一步中的用

户对象推荐的物品集(本文实验推荐电影集);

第三步 计算采用某种差分隐私保护处理后给第一步中的用户对象推荐的物品集;

第四步 求两个推荐物品集的交集;

第五步 用该交集除以推荐物品集的个数,计算出一个百分比,该百分比值越大意味着 ε 取值对推荐的准确率影响越小,此时的 ε 选择就应该相对合理。

需要说明的是:该方案仅能给出一个比较合理的 ε 的取值范围。如果由上述第五步产生的百分比小于 20%,则认为虽然隐私保护力度很强,但是已经严重影响了推荐结果;当百分比大于 80%,则认为虽然有了良好的推荐结果,但隐私保护力度过弱。因此,为了权衡隐私保护力度和推荐准确率,认为该百分比在 20%～80%之间的 ε 取值均是比较合理的选择,即一方面实现了差分隐私保护,另一方面也没有严重影响推荐结果。

限于篇幅,在此仅给出用于 ALS 目标函数加扰算法中的 ε 选择情况及结果。实验中,关于 ALS 的参数仍按照4.3.1节设置。此外,实验推荐电影集的个数设为 20,随机任选一个用户作为推荐的对象,数据集为 Movielens-1M。由于噪声的随机性,结果取 10 次的平均值。图 5 示出了 ε 取值对推荐准确率的影响。

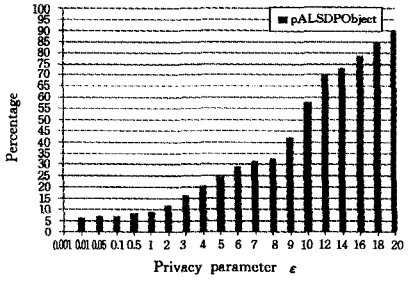


图 5 隐私保护参数 ε 对推荐结果的影响

从图 5 可以看出,当 ε 取值为[4,16]时,未做差分隐私保护和做了差分隐私保护的算法分别得到的推荐电影集的重合度百分比取值为[20%,80%],即该百分比范围中的 ε 取值均是可接受的。

结束语 无论从学术还是商业价值的角度,推荐系统的绝大多数研究都是考虑如何提高推荐的准确率,而忽略了用来做研究的原始数据的隐私保护问题。譬如,用户仅仅从表面上看问题,以为自己为电影评分、留下商品的购买记录和浏览记录等不会直接泄露个人隐私,实际上,如果攻击者具有一定的背景知识,那么个人隐私是完全可以被间接推导出来的。文献[1]将当今被证明为最为严格的差分隐私保护技术应用于解决推荐问题的矩阵分解之中,但存在缺少满足差分隐私保护的证明等问题。本文一方面补充了相应的证明,同时也指出某些算法中的错误;另一方面还将 Chaudhuri 等人提出的目标函数加扰方法灵活应用于 ALS 中。最后通过在两个真实数据集上的实验验证结果表明本文的算法 DPALSObject 获得了更优的推荐准确率,且推荐准确率不会随差分隐私保护参数发生显著的变化,即趋于平缓。

目前随着 Internet 的飞速发展和越来越多的用户期望更个性化的推荐服务,隐私保护问题也必将成为用户顾虑的问题,推荐系统乃至数据挖掘领域要健康地发展离不开对隐私保护问题的深入研究。今后可从以下几个方面来展开更深入的研究:

1)本文对隐私保护参数 ε 的合理选择仅仅给出了一个范围,未能给出权衡隐私保护和推荐结果到最优的方案。

2)本文算法为了取得较好的推荐结果,隐私保护参数 ε 的取值都偏大,因此可以继续改进算法,以实现用更小的 ε 获得更高的推荐准确率。

3)算法中其余参数的自动调节问题。

参 考 文 献

[1] BERLIOZ A, FRIEDMAN A, KAAFAR M A, et al. Applying Differential Privacy to Matrix Factorization[C]// Proceedings of the 9th ACM Conference on Recommender Systems. Vienna, Austria, 2015: 107-114.

[2] DWORK C. Differential Privacy[C]// 33rd International Colloquium on Automata, Languages and Programming, part II (ICALP 2006). Italy, 2006: 1-12.

[3] CHAUDHURI K, MONTELEONI C, SARWATE A. Differentially private empirical risk minimization[J]. Journal of Machine Learning Research, 2011, 12: 1069-1109.

[4] DESROSIERS C, KARYPIS G. A comprehensive survey of neighborhood-based recommeddation methods [M] // Rec. Sys. Handbook. 2001: 107-144.

[5] CANNY J. Collabrative filting with privacy[J]. IEEE Security and Privacy, 2002, 18(1): 45-57.

[6] POLAT H, DU W. Achieving private recommendations using randomized response techniques[C]// PAKDD. 2006: 637-646.

[7] NIKOLAENKO V, IOANNIDIS S, WEINSBERG U, et al. Privacy-preserving matrix factorization[C]// CCS. 2013: 801-812.

[8] CALANDRINO J A, KILZER A, NARAYANAN A, et al. "you might also like:" privacy risks of collaborative filtering[C]// IEEE Security and Privacy. 2011: 231-246.

[9] CHEN R, MOHAMMED N, FUNG B C M, et al. Publishing Set-Valued Data via Differential Privacy[C]// Proceedings of the 37nd Conference of Very Large Databases (VLDB). Seattle, Washington, USA, 2011: 1087-1098.

[10] CHEN R, FUNG B C M, et al. Differentially Private Transit Data Publication: A Case Study on the Montreal Transportation System[C]// Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD). Beijing, China, 2012: 493-502.

[11] FRIEDMAN A, SCHUSTER A. Data Mining with Differential Privacy[C]// Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD). Washington, DC, USA, 2010: 493-502.

[12] MACHANAVAJJHALA A, KOROLOVA A, SARMA A D. Personalized social recommendations-arrcurate or private[J]. PVLDB, 2011, 4(7): 440-450.

[13] MCSHERRY F, MIRONOV I. Differential Private recommender System: Building privacy into Netflix prize conteders[C]// KDD. Paris, France, 2009: 627-636.

(下转第 94 页)

dology And Software Platform For Probabilistic Risk Assessment And Safety Monitoring Of Socio-Technical Systems[J]. Reliability Engineering & System Safety, 2010, 95(12): 1276-1285.

[4] CUI L, SHU Y D, WANG Z H, et al. HASILT: An Intelligent Software Platform for HAZOP, LOPA, SRS and SIL Verification[J]. Reliability Engineering & System Safety, 2012, 108: 56-64.

[5] XU X Z, WANG G Y, DING S F, et al. A New Method For Constructing Granular Neural Networks Based On Rule Extraction And Extreme Learning Machine[J]. Pattern Recognition Letters, 2015, 67: 138-144.

[6] DARAAMOLA O, STALHANE T, MOSER T, et al. A Conceptual Framework for Semantic Case-based Safety Analysis[C]// Emerging Technologies & Factory Automation (ETFA). IEEE, 2011: 1-8.

[7] ZHANG Y J, SUN Y C. Safety Risk Assessment of Human-machine Interaction Behavior in Cockpit[C]// Computational Intelligence and Design (ISCID). IEEE, 2014: 39-42.

[8] ZHOU H M, HUANG G B, LIN Z P, et al. Stacked Extreme Learning Machines [J]. IEEE Transactions on Cybernetics, 2015, 45(9): 2013-2025.

[9] LIANG N Y, HUANG G B, SARATCHADRAN P, et al. A Fast and Accurate Online Sequential Learning Algorithm for Feedforward Networks[J]. IEEE Transactions on Neural Networks, 2006, 17(6): 1411-1423.

[10] JAVED K, GOURIVEAU R, ZERHOUNI N. A New Multivariate Approach for Prognostics Based on Extreme Learning Machine and Fuzzy Clustering[J]. IEEE Transactions on Cybernetics, 2015, 45(12): 2626-2639.

[11] ZHANG N, DING S F, SHI Z Z. Denoising Laplacian Multi-Layer Extreme Learning Machine[J]. Neurocomputing, 2016, 171: 1066-1074.

[12] MCDONNELL M D, TISSERA M D, VLADUSICH T, et al. Fast, Simple and Accurate Handwritten Digit Classification by Training Shallow Neural Network Classifiers with the 'Extreme Learning Machine' Algorithm[J]. Plos One, 2015, 10(8).

[13] DENG W Y, ZHENG Q H, CHEN L, et al. Research On Extreme Learning of Neural Networks [J]. Chinese Journal of Computers, 2010, 33(2): 279-287. (in Chinese)
邓万宇, 郑庆华, 陈琳, 等. 神经网络急速学习方法研究[J]. 计算机学报, 2010, 33(2): 279-287.

[14] BUENO-CRESPO A, GARCIA-LAENCINA P J, SANCHO-GOMEZ J L, et al. Neural Architecture Design Based On Extreme Learning Machine[J]. Neural Networks, 2013, 48: 19-24.

[15] JOLLIFFE I T. Principal Component Analysis[M]. New York: Wiley Online Library, 2010.

[16] MOZAFFARI A, LASHGARIAN A N, FATHI A. Regularized Machine Learning Through Constraint Swarm And Evolutionary Computation Applied To Regression Problems[J]. International Journal of Intelligent Computing and Cybernetics, 2014, 7(4): 346-381.

[17] ZHU W T, MIAO J, QING L Y. Constrained Extreme Learning Machine: a Novel Highly Discriminative Random Feedforward Neural Network[C] // Proceedings of The 2014 International Joint Conference On Neural Networks. IEEE, 2014: 800-807.

[18] ZHAO R, MAO K Z. Semi-Random Projection for Dimensionality Reduction and Extreme Learning Machine in High-Dimensional Space [J]. IEEE Computational Intelligence Magazine, 2015, 10(3): 30-41.

[19] LI C, RENE-VINICIO S, ZURITA G, et al. Multimodal Deep Support Vector Classification With Homologous Features And Its Application To Gearbox Fault Diagnosis[J]. Neurocomputing, 2015, 168: 119-27.

[20] LIU B, XIA S X, MENG F R, et al. Manifold regularized extreme learning machine [J]. Neural Computing and Applications, 2016, 27(2): 255-269.

(上接第 88 页)

[14] LIU Z Q, WANG Y X, SMOLA, et al. Fast Differentially Private Matrix Factorization[C] // Proceedings of the 9th ACM Conference on Recommender Systems. Vienna, Austria, 2015: 171-178.

[15] HUA J Y, XIA C, ZHONG S. Differentially private matrix factorization[C]// Proceedings of the 24th International Conference on Artificial Intelligence. Buenos Aires, Argentina, 2015: 1763-1770.

[16] SWEENEY L. K-anonymity: A model of for protecting privacy [J]. International Journal of Uncertainty, Fuzziness and Knowledge Based System, 2002, 10(5): 557-570.

[17] LI Y, ZHANG X Z. Survey of Research on Differential Privacy [J]. Journal of Application Research of Computers, 2012, 29 (9): 3201-3211. (in Chinese)
李杨, 张新政. 差分隐私保护综述[J]. 计算机应用研究, 2012, 29 (9): 3201-3211.

[18] ZHANG X J, MENG X F. A Survey of Differential Privacy in Data Publication and Analysis[J]. Chinese Journal of Computers, 2014, 37(4): 929-949. (in Chinese)
张啸剑, 孟小峰. 面向数据发布和分析的差分隐私保护研究[J]. 计算机学报, 2014, 37(4): 929-949.

[19] DWORK C. Differential Privacy: A Survey of Results[C]// Theory and Applications of Models of Computation (TAMC2008). Berlin, 2008: 1-9.

[20] NISSIM K, RASKHODNIKOVA S, SMITH A. Smooth sensitivity and sampling in private data analysis[C]// Proceedings of the 39th Annual ACM Symposium on Theroy of Computing. San Diego, USA, 2007: 75-84.

[21] DWORK C, MCSHERRY F, NISSIMK, et al. Calibrating Noise to Sensitivity in Private Data Analysis[C]// Proceedings of the 3th Theory of Cryptography Conference (TCC). New York, USA, 2006: 363-385.

[22] MCSHERRY F. Privacy Int egrated Queries: An Extensible Platform for Privacy-Preserving Data Analysis[C]// Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD). Providence, Rhode Island, USA, 2009: 19-30.