联邦推荐系统综述

朱智韬[1,2]，司世景[1]，王健宗[1]，肖京[1]
1. 平安科技（深圳）有限公司，广东 深圳 518063；
2. 中国科学技术大学，安徽 合肥 230026

摘要
在联邦学习范式中，原始数据被本地存储在独立的用户客户端中，而脱敏数据被发送到中心服务器中加以聚合，这给众多领域提供了一种新颖的设计思路。考虑到传统推荐系统的研究方向集中于提高推荐效果，在资源节约、跨领域推荐、隐私保护等方面还具有很大改进空间，如何将联邦学习与推荐系统结合以解决这些问题成为当前的一个研究热点。对近年来基于联邦学习的推荐系统进行了全面的总结、比较与分析，首先介绍了推荐系统的传统实现方式及面临的瓶颈；然后引入了联邦学习范式，描述了联邦学习在隐私保护、利用多领域用户数据两方面给推荐系统带来的增益，以及二者结合的技术挑战，进而详细说明了现有的联邦推荐系统部署方式；最后，对联邦推荐系统未来的研究进行了展望与总结。

关键词
联邦学习；推荐系统；隐私保护；协同过滤；深度学习

中图分类号：TP391 文献标志码：A doi: 10.11959/j.issn.2096-0271.2022032

Survey on federated recommendation systems

ZHU Zhitao[1,2], SI Shijing[1], WANG Jianzong[1], XIAO Jing[1]
1. Ping An Technology (Shenzhen) Co., Ltd., Shenzhen 518063, China
2. University of Science and Technology of China, Hefei 230026, China

Abstract
In the federated learning (FL) paradigm, the original data are stored in independent clients while masked data are sent to a central server to be aggregated, which proposes a novel design approach to numerous domains. Given the wide application of recommendation systems (RS) in diverse domains, combining RS with FL techniques has been gaining momentum to reduce the computational cost, do cross-domain recommendation and protect users' privacy while maintaining recommendations performance as traditional RS. The federated learning-based recommendation systems in recent years were comprehensively summarized. The difference between traditional and federated recommendation systems was analyzed, and the main research direction and progress of federated recommendation systems were demonstrated with comparison and analysis. Firstly, the traditional recommendation systems and their bottleneck were summarized. Then the federated learning paradigm was introduced. Furthermore, the advantages of combining federated learning with

recommendation systems were depicted in two aspects: privacy protection and usage of multi-domain user information, along with the technical challenges during the combination. At the same time, the existing deployment of federated recommendation systems was illustrated in detail. Finally, future research on federated recommendation systems was prospected and summarized.

Key words
federated learning, recommendation system, privacy-preserving, collaborative filtering, deep learning

0 引言

伴随着互联网的快速发展，信息过载问题成为每个人的生活中一个愈发严重的阻碍。作为解决信息过载问题的有效解决方案，推荐系统（recommendation system，RS）在商业网站和信息分发应用中普遍存在。它们利用用户的各种知识和数据生成个性化推荐结果，是抵御客户无意义浏览和商家无用数据推送的利器。一般来说，推荐列表立足于格式化存储的三方面信息：用户喜好、项目属性、用户与项目的交互记录。此外，还可能使用其他附加信息，如时间和空间数[1]。推荐系统服务的最终目标是提高营业额[2]，根据不同的应用场景，服务提供者可能会采取不同的路径来实现这一目标，包括向目标用户推荐合适商品以直接增加销售量，以及推送匹配用户兴趣的娱乐内容以提高用户黏度，从而使广告获得更多的曝光率等。但毫无疑问，各种隐私保护法规的颁布对跟踪用户足迹的程序的要求将变得越来越严格[3]。越来越多的大企业开始意识到隐私保护和数据安全合规的必要性，并出现了基于联邦学习（federated learning，FL）等机制的隐私保护措施。

本文对推荐系统、联邦学习以及两者的结合进行了概述。首先阐述推荐系统的作用机制，然后介绍联邦学习的发展源流，最终的目标是在不破坏数据隐私和遵守安全原则的前提下，探索最大限度地利用数据的联邦推荐方法。

1 推荐系统

在实践中，推荐系统已经被广泛应用于书籍和CD[4]、音乐[5]、电影[6]、新闻[7]、笑话[8]和网页[9]等推荐中。根据推荐机制的不同，传统推荐模型主要分为协同过滤式推荐系统、基于内容的推荐系统和混合推荐系统[10]。而后深度学习方法的兴起为推荐系统的研究带来了新的机遇，本文也将举例说明其结合思想。

1.1 协同过滤推荐系统

协同过滤算法的核心思想是根据相似用户或相似物品来提供商品推荐或预测，其包括两个维度：基于邻域和基于模型。最初的邻域算法集中在用户之间的相似性上，被称为基于用户的协同过滤（user-based collaborative filtering，UCF）；基于项目的协同过滤（item-based collaborative filtering，ICF）方法[11]则关注类似项目的评分。基于邻域的协同过滤推荐的关键是计算不同用户、不同项目之间的相似度。**表1**给出了比较常用的相似度度量方法。

基于邻域的协同过滤系统过去非常成

表1 相似度度量方法

相似度度量	表达式	优点	缺点
闵可夫斯基距离	$\mathrm{dist}(X,Y)=\left(\sum_{i=1}^{n}\lvert x_i-y_i\rvert^p\right)^{\frac{1}{p}}$	反映个别数值特征的绝对差异	受离群值影响较大
杰卡德相似系数	$J(A,B)=\frac{\lvert A\cap B\rvert}{\lvert A\cup B\rvert}$	计算简单，算力负担小	没有考虑用户之间的评分差异；精度不高
余弦相似度	$\cos(\theta)=\frac{\boldsymbol{a}^{\top}\boldsymbol{b}}{\lvert\boldsymbol{a}\rvert\cdot\lvert\boldsymbol{b}\rvert}$	对绝对值不敏感，修正了用户之间可能存在的参数差异	将缺失值设置为0会缩小真实差距
皮尔逊相关系数	$\rho_{XY}=\frac{\sum_{i=1}^{n}(X-\mu_X)(Y-\mu_Y)}{\sqrt{\sum_{i=1}^{n}(X-\mu_X)^2}\sqrt{\sum_{i=1}^{n}(Y-\mu_Y)^2}}$	用平均值填充缺失的值，以避免分数膨胀	不能直接反映所使用的物品数量；在稀疏的数据集上不易计算相似度

功，得到了广泛使用，但受到可扩展性不足（最近邻算法需要的计算量随着用户数和项目数的增长而快速增长，不适合应用于数据量较大的场景）的限制。因此，有学者提出了一种基于模型的协同过滤（model-based collaborative filtering，MCF）算法，利用机器学习或数据挖掘等算法，通过训练数据对复杂模式进行学习和识别，得到学习模型，然后根据学习模型对数据集进行智能预测[12-13]。常用的模型协同过滤算法有隐语义协同过滤模型[11]、隐因子模型等。

此外，还有其他路径的协同过滤推荐算法，如将推荐问题转化为节点选择问题的图模型推荐算法[14]、模拟人类推理因果关系的不定性的贝叶斯网络协同过滤模型[15]、聚焦用户属性特征的聚类协同过滤模型[16]等。

作为目前使用非常广泛的推荐算法，协同过滤推荐系统工程实现简单，模型通用性强，效果显著，但始终面临着严重的数据稀疏（用户仅对数据库中可用项目的极少部分进行评分）与冷启动（新用户与新物品缺少评分数据作为可学习的历史信息）问题，并且通常使用的浅层模型无法学习到用户和项目的深层次特征[17]。

1.2 基于内容的推荐系统

基于内容的推荐系统（content-based recommendation system，CB）算法通过寻找与用户已选择项目具有相似属性的物品进行推荐。Pazzani M J等人[18]对基于内容的推荐策略架构进行了宏观介绍，这种推荐只需要两类信息，即对项目特征的描述和用户过去的偏好信息，不需要大量的用户评分历史来生成推荐列表（避免了评分数据稀疏问题）。实现的关键点在于对项目和用户偏好进行建模，并计算其相似性（此时提取新项目的特征即可进行冷启动）。Salton G等人[19]提出的向量空间模型是非常常用的内容建模方法。基于内容的推荐系统可解释性强，可以更好地解决冷启动问题，对于小众领域也有较好的推荐效果，其缺点在于推荐精准度较低、无法挖掘用户深层次的潜在兴趣，导致推荐新颖度不高、较难分发长尾标的物，并且这种方法的有效性与可扩展性受到人工设计特征提取方法的严重制约，常常会遇到特

征提取困难的问题。

1.3 混合推荐系统

类似于集成学习通过有效整合不同算法降低系统性误差的思想，混合推荐系统搭配使用不同的推荐算法给出最终的推荐结果，避免单一算法固有的问题，实现较任意单一算法更佳的推荐效果[20]。此外，一些混合推荐算法融合了包括图像、音频、文本在内的多源异质辅助信息，能够有效解决数据稀疏和冷启动问题。但囿于辅助信息的复杂性（如异质性、多模态、数据提供方限制等），相关方法的研究仍然面临着严峻挑战[21]。

值得一提的是，基于其他方法的推荐系统也各有千秋，如基于社交网络的[22]、基于人口数据的[23]、基于心理学的[24]、基于大数据的，等等。每一种推荐算法都有其优缺点，本文篇幅有限，不做深入阐述。

1.4 基于深度学习的推荐系统

随着算力的跨越式提高，深度学习已经成为互联网人工智能的一个利器。神经体系结构在有监督和无监督学习任务中都获得了巨大的成功[25-26]。自然地，深度学习也在推荐系统领域得到了广泛应用，并且得益于其在非线性转换、表征学习、序列模型以及灵活性方面的优势，成为当前的最优模型[1]。

深度学习在推荐系统中的应用最早可以追溯到2007年Salakhutdinov R等人[27]发表的一篇将受限玻尔兹曼机应用于推荐系统的文章。常用的深度学习模型有多层感知机、卷积神经网络（convolutional neural network，CNN）、循环神经网络（recurrent neural network，RNN）、注意力模型、自编码器、神经自回归分布估计、对抗网络、受限玻尔兹曼机、深度强化学习等，这些模型在推荐系统上得到了广泛的应用[1]。

基于深度学习的推荐系统通常将与各类用户和项目相关的数据作为输入，利用深度学习模型学习用户和项目的隐表示，并基于这种隐表示为用户生成项目推荐[17]。目前构建深度学习推荐算法最常见的一种范式是多层感知机，其可将非线性变换添加到现有的推荐系统方法中，如果需要整合附加信息（图像、文本、语音、视频等），则会采用CNN、RNN模型来提取相关信息。

而后根据所采用的神经网络的不同结构，演变出了不同的深度学习推荐模型。

通过增加深度神经网络结构的层数和复杂度，AutoRec引入了自编码器，如**图1**所示，利用单隐层神经网络模型，结合协同过滤的共现矩阵，得到用户/物品向量的自编码，后续生成用户对物品的预估评分并用于排序[28]。

通过丰富特征交叉方式演变出了神经协同过滤（neural collaborative filtering，NCF）[29]，**图2**中摈弃了矩阵分解中的简单内积操作，“多层神经网络+输出层”让用户/物品向量进行更充分的交叉，引入更多的非线性特征，增强对稀疏特征的学习能力。

通过引入注意力机制，在嵌入层与多层感知机之间加入注意力层，演变出了深度兴趣网络（deep interest network，DIN）[30]，以及融合了序列模型以模拟用户喜好变化过程的深度兴趣进化网络（deep interest evolution network，DIEN）[31]、使用胶囊网络提取用户的多样兴趣并引入基于标签的注意力机制的动态路径选择多兴趣网络（multi-interest network with dynamic routing，MIND）[32]等。

而对于因子分解机模型的各种深度学习演化，克服了协同过滤对稀疏矩阵泛

化能力不强的困难，包括神经因子分解机（neural factorization machine，NFM）应用神经网络改善因子分解机二阶交叉部分的特征交叉性能[33]、因子分解机神经网络（factorization-machine supported neural network，FNN）利用因子分解机的结果初始化网络[34]、注意力神经因子分解机（attention neural factorization machine，AFM）在双线性交互池化操作中引入注意力机制以提升模型的表示能力与可解释性[35]。

对于图像、文本、音乐数据推荐任务，通过将CNN作为特征处理手段并应用到推荐系统中，可以实现对多种类推荐项目的处理，如基于文本与图像数据推荐微博标签[36-37]、利用深度学习模型解决音乐推荐场景下的冷启动问题[38]、使用比较深度学习（comparative deep learning，CDL）方法将用户和图像映射到同一隐空间中以推荐图像[39]。

针对序列数据，一般为了捕捉用户行为间的相互依赖关系，可以使用RNN建模实现项目推荐和用户行为预测. 可以利用的数据类型包括历史会话行为记录[40]、用户行为的时间序列信息[41-42]、注意力机制下的文本序列特征[43]等。

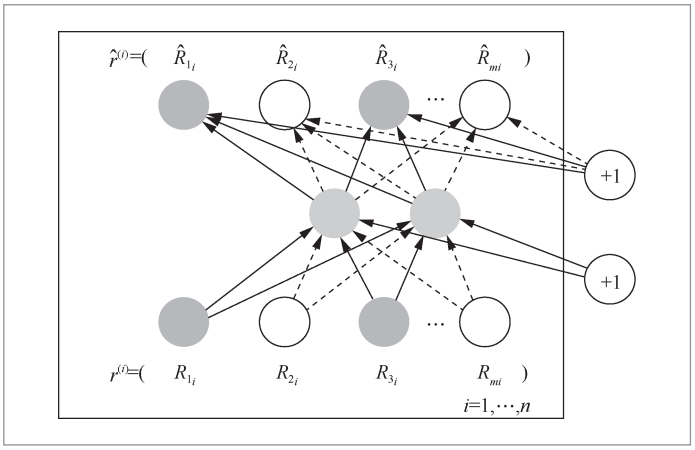


图1 自编码器深度推荐系统

2 联邦推荐系统

2.1 联邦学习

与许多机器学习算法一样，推荐系统

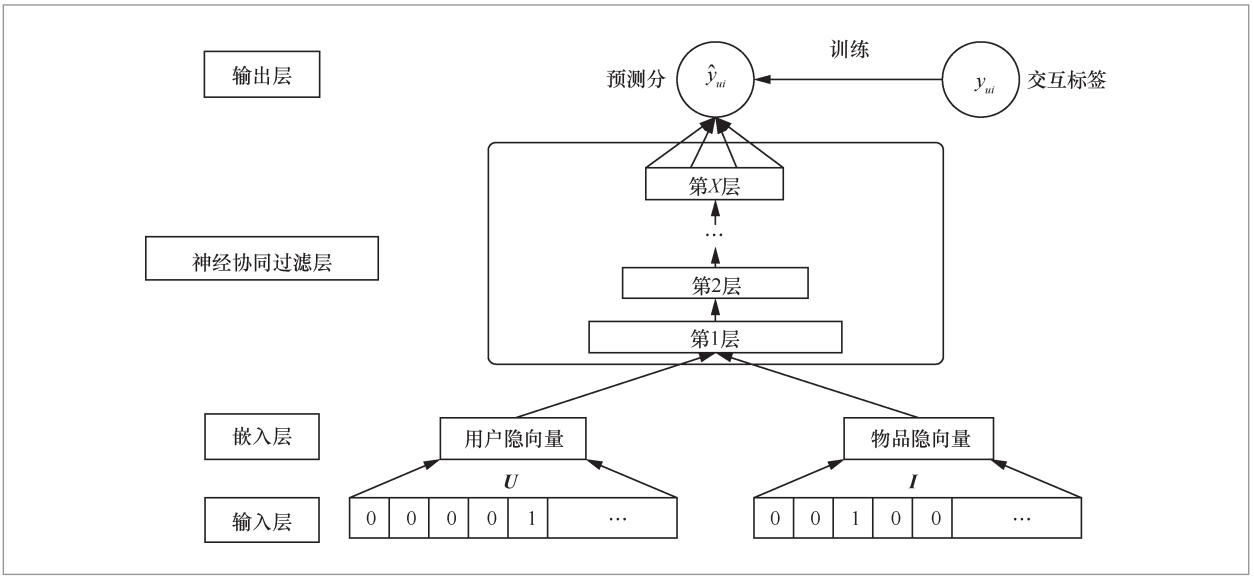


图2 神经协同过滤

也是“数据饥渴”的，但现实情况是，数据被置于不同组织的隐私保护下，且受到各种政策法规的限制，例如欧盟的《通用数据保护条例》（General Data Protection Regulation，GDPR）[3]。数据不能简单地在机构之间共享，不同公司和组织间成为“数据孤岛”。由于每个“数据孤岛”中数据的大小或特征都有局限性，因此单个机构可能无法训练出一个高质量的推荐模型。

联邦学习是一种机器学习环境，由McMahan B等人[44-45]首次提出，Kairouz E B P等人[46]给出其严谨的定义：联邦学习是一种机器学习环境，在中央服务器或服务提供商的协调下，多个实体（客户端）协作解决机器学习问题。每个客户端的原始数据都存储在本地，不进行交换或传输；相反，旨在进行即时聚合的集中更新被用来实现学习目标。对于一个有效的联邦学习系统，只要求存在任何一个客户端能从联邦学习系统中获得更高的模型效用[47]。

在联邦学习框架中，学习任务是由松散的参与方联邦（本文称之为客户端）共同完成的，如**图3**所示，原始数据被各自独立存储在本地，这些设备由一个中央服务器协调。最早的实践项目甚至在数千万部手机和边缘设备应用中部署了联邦学习[48]。

按照客户端是单一设备还是组织或公司，联邦学习可以分为跨设备联邦学习和跨孤岛联邦学习。而根据在不同参与方之间的训练数据特征空间和样本ID空间的分布情况，Yang Q等人[49]将联邦学习划分为横向联邦学习、纵向联邦学习以及联邦迁移学习。

2.1.1 横向联邦学习

横向联邦学习也可被称为基于样本的联邦学习，它是在不同数据集之间特征空间重叠较多而用户重叠较少的场景下引入的。

这种联邦学习对金融机构最有吸引力。因为金融机构的用户群体通常被限制

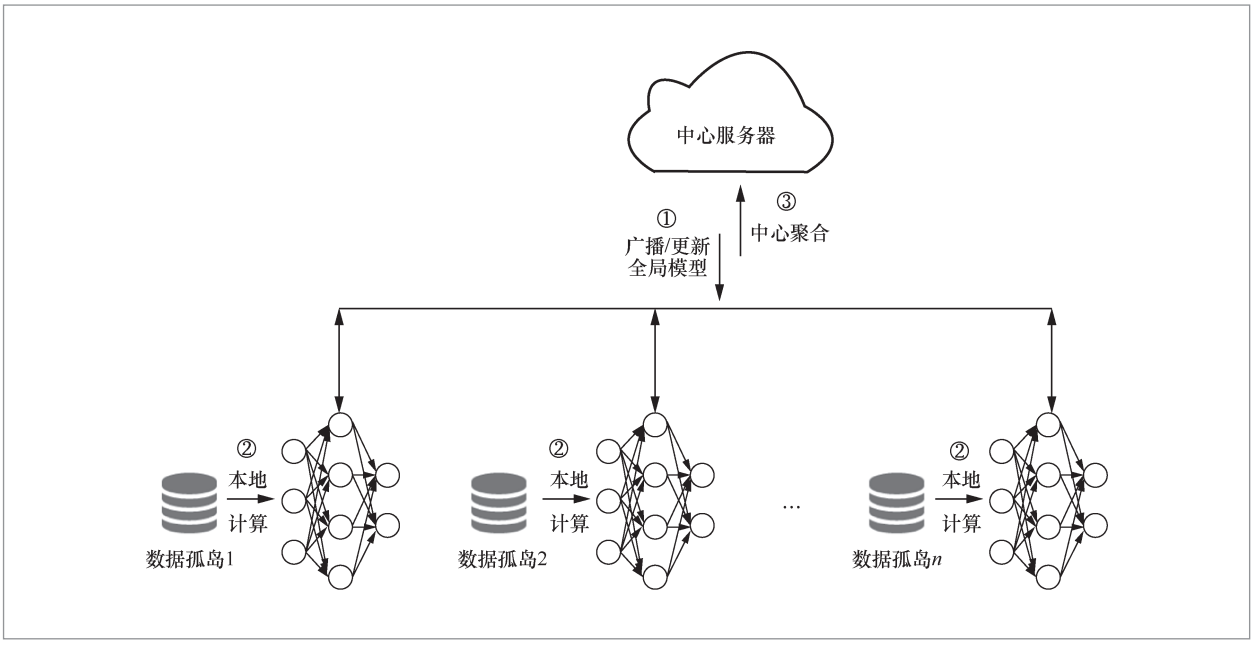


图 3 联邦学习框架

在各自的区域，业务种类相差不多，因此可以从用户维度对数据集进行划分，使用特征相同而用户不完全相同的双方部分数据共同训练。

2.1.2 纵向联邦学习

纵向联邦学习也叫基于特征的联邦学习，适用于数据持有者之间存在重叠的数据样本但数据特征不同的情形。纵向联邦学习的客户端通常由同区域不同业务的公司组成，其用户集包含了该地区的大部分居民，但错开了购买特征。应用纵向联邦学习可以拓宽双方的特征空间，预测最值得推销的商品。

虽然纵向联邦学习可以丰富用户特征，发现不同数据之间的隐藏关系，但不同数据特征的整合通常更具挑战性[49]。

2.1.3 联邦迁移学习

联邦迁移学习考虑了数据方在用户空间或特征空间中仅有部分重叠的挑战性场景，不对数据进行切分，而是利用现有的迁移学习技术[50]协作建立模型，以克服样本或标签不足的困难。

令第 i 方数据 $\mathcal{D}_i$ 的样本空间为 $\mathcal{X}_i$，特征空间为 $\mathcal{Y}_i$，标签为 I_i，则联邦学习的3种分类表示见**表2**。

显然，推荐系统的工作方式在个性化推荐收益与隐私保护之间形成了一种权衡。随着公众对大公司收集和使用个人数据的举措越来越感到不安，推荐系统如何在个性化和隐私保护之间寻求最优解变得愈发重要。另外，考虑到计算成本和数据的可扩展性，在推荐系统上应用联邦学习的尝试也是相当明智的。

当前联邦推荐系统的结合出发点多为增强隐私保护与促进多领域用户信息融合，而实际落地阶段往往面临数据异质性与通信成本增加的挑战。

表 2 联邦学习分类

分类	表示
横向联邦学习	$\mathcal{X}_i=\mathcal{X}_j,\mathcal{Y}_i=\mathcal{Y}_j,I_i\neq I_j \quad \forall\mathcal{D}_i,\mathcal{D}_j,i\neq j$
纵向联邦学习	$\mathcal{X}_i\neq\mathcal{X}_j,\mathcal{Y}_i\neq\mathcal{Y}_j,I_i=I_j \quad \forall\mathcal{D}_i,\mathcal{D}_j,i\neq j$
联邦迁移学习	$\mathcal{X}_i\neq\mathcal{X}_j,\mathcal{Y}_i\neq\mathcal{Y}_j,I_i\neq I_j \quad \forall\mathcal{D}_i,\mathcal{D}_j,i\neq j$

2.2 联邦推荐系统的隐私保护手段

推荐系统的高性能建立在海量的数据挖掘基础之上，一方面，为了达到足够的可用性，推荐系统对用户历史数据的体量和具体程度有着严苛的要求；另一方面，用户的历史数据提供得越详实，其隐私信息泄露的风险越大。因此，如何解决推荐系统中的隐私保护问题，即在保护用户历史行为数据、推荐模型与推荐结果等隐私信息的前提下，达到推荐结果的可靠性与有效性等功能指标，是一个亟待解决且非常具有理论意义的问题。常用的隐私保护方法大致分为以下两个类别。

- 密码学方法。将输入数据隔绝于其他参与方或者不以明文传输的方式，使分布式计算过程不泄露隐私信息，如安全多方计算（包括不经意传输、秘密共享、混淆电路和同态加密）[51]。
- 模糊处理。随机化、添加噪声或修改数据使传输数据拥有某一特定级别的隐私，如差分隐私方法。

传统基于内容的推荐系统隐私保护系统架构利用公钥全同态加密等技术，先在本地对用户历史数据训练集中的输入数据进行加密，而后由推荐服务器在密文域上建立预测模型，计算推荐结果，接着返回给授权用户，以解密推荐结果并验证其正

确性。

而基于邻域的推荐系统隐私保护系统架构则先将与目标用户组相似用户的历史数据作为数据集，各推荐服务器在密文域上，通过安全多方计算等技术建立预测模型并计算推荐结果，并将安全多方计算得到的密文形式推荐结果与其正确性可验证证据共同返回给目标用户组的推荐服务器，随后分发给目标用户以解密验证结果。现有的做法通常是利用公钥全同态加密技术与混淆电路技术，在假定不存在合谋攻击的推荐服务器与密码服务提供商间通过安全多方计算实现[52]。Kim S等人[53]使用全同态加密技术部署隐私保护的矩阵分解推荐系统，针对加密后的用户评级数据输入，实施矩阵分解并输出加密结果。

上述工作均使用公钥（全）同态加密技术实现，虽然模型较为准确，但计算开销和密文长度随着用户数据集的扩大而急剧增大，给资源受限的用户端带来了难以承受的复杂度与通信开销。并且，基于公钥全同态加密技术与混淆电路的基于邻域的隐私保护协同过滤推荐系统更加难以抵御半可信或恶意环境下的推荐服务器攻击[52]。

在保证推荐效果近似可用性的前提下，数据扰动技术已经被广泛应用于推荐系统。Agrawal R等人[54]首次将加法扰动技术应用于数据挖掘领域，并验证了其在决策树等算法上的效果。通过在原始数据矩阵后加上一个扰动矩阵（各行由一个均值为0的均匀分布或高斯分布独立生成），达到使原始数据失真并且不改变计算结果统计均值的目的。设原始向量为 $\boldsymbol{A}$、$\boldsymbol{B}$，扰动向量分别为 $\boldsymbol{R}$、$\boldsymbol{V}$，则 $\boldsymbol{A}$、$\boldsymbol{B}$ 的内积可由 $\boldsymbol{A'}=\boldsymbol{A}+\boldsymbol{R}$、$\boldsymbol{B'}=\boldsymbol{B}+\boldsymbol{V}$ 的内积近似表示：

$$\boldsymbol{A'}\cdot\boldsymbol{B'}=\sum_{i=1}^{n}(a_i+r_i)(b_i+v_i)=\sum_{i=1}^{n}(a_ib_i+a_iv_i+r_ib_i+r_iv_i)\approx\sum_{i=1}^{n}a_ib_i \tag{1}$$

在此基础上，Polat H等人[55]与Herlocker J等人[56]分别基于均匀分布扰动因子与高斯分布扰动因子构建了保护隐私的推荐系统。而后Chen K K等人[57]利用隐私保护的欧几里得距离计算与隐私保护的内积计算技术，构建了基于乘法扰动的二分类隐私保护推荐系统。

基于前人的工作，Dwork C[58]最早提出差分隐私（differential privacy，DP）的概念，提供了个人隐私泄露的数学定义，并根据应用场景的不同将差分隐私分为中心化差分隐私和本地化差分隐私，前者由可信的数据收集者添加噪声，后者由用户本地添加噪声。中心化差分隐私的典型噪声机制是拉普拉斯噪声机制和指数噪声机制[57]，其中前者用来处理连续型数据，后者适用于离散型数据。本地化差分隐私则主要采取随机响应方法来保护隐私[59]。

在差分隐私的定义下，攻击者无法通过查询结果的改变探知具体数据的内容，并且数据集依然保持可用于数据挖掘等操作的一些统计特性。虽然差分隐私技术基于数据扰乱添加噪声，使得原有的数据失真，但其加入的噪声量大小只与数据集的敏感度和隐私参数 ε 有关，而与数据集大小无关，因此对于大规模的数据集仍然可能不提出过高的性能要求。这使得差分隐私保护技术可以在确保数据可用性的前提下大幅度降低隐私泄露风险。其缺点在于，在复杂计算过程中，噪声累积有可能导致数据不可用。

为了保护推荐系统的差分隐私，McSherry F等人[60]将差分隐私技术应用于协同过滤算法，通过在用户评分数据的计数与求和中添加采用拉普拉斯机制计算的噪声，得到了盲化的物品-物品协方差矩阵，实现了差分隐私保护的K近邻（K-nearest neighbor，KNN）与奇异

值分解(singular value decomposition, SVD)推荐算法。Zhu T Q等人[61]则对基于邻域的协同过滤推荐算法的邻居选择环节进行差分隐私保护,并对预测的评分值进一步添加拉普拉斯噪声以保护预测结果。Berlioz A等人[62]则针对矩阵分解算法的3个环节(数据输入、矩阵分解进程、数据输出)分别应用差分隐私技术,并权衡了各方法在隐私保护力度与推荐准确度之间的取舍,如**图4**所示。

对于严格联邦定义下的推荐系统,最初的联邦学习认为,由于其中心服务器只从参与方采集梯度更新信息,这种分布式的模型更新方式已经能在一定程度上保护隐私,但Orekondy T等人[63]与Wang Z B等人[64]证明了没有隐私保护机制的简单联邦学习仍会遭受梯度信息反推攻击,从而泄露隐私信息;Melis L等人[65]对参与人数较少的协同学习下非预期特征泄露情景进行了成员推理攻击与属性推理攻击,并提出了多种防御建议。

受到诸多隐私保护技术的启发,联邦学习广泛运用了安全多方计算、同态加密以及差分隐私等隐私保护技术[66],但是其在应用场景上与传统推荐系统的隐私保护方法有很大的不同。传统推荐系统的隐私保护致力于防止推荐服务提供商对单一客户端原始用户数据的探知,通过对推荐模型的原始输入进行加密传输,存储、计算资源受限的用户可以将推荐算法外包给资源庞大的推荐服务器;而联邦推荐系统的隐私保护则是针对不同的推荐服务提供商实体,在不交换各自本地数据的前提下,对推荐模型的梯度更新进行加密传输,以分摊计算开销,并且规避直接在原始数据密文域上训练全局模型的庞大计算开销,有效地降低了传统机器学习源数据聚合带来的许多隐私风险[67]。相比于传统的一方提供数据、另一方提供计算服务的推荐范式,联邦学习使得原先的用户也成为潜在的推荐服务提供商,扩宽了推荐数据的来源,对于单一客户端而言可以有效提高推荐准确性。

Truex S等人[68]开发了一种联邦训练方法,可以根据本地隐私预算对复杂模型参数更新执行基于本地化差分隐私(local differential privacy, LDP)机制的扰动,同时最大限度地降低噪声对联邦学习训练过程的影响,验证了使用压缩LDP协议部署的系统针对公共数据训练深度神经网络时的有效性。Liu R X等人[69]提出了一个用

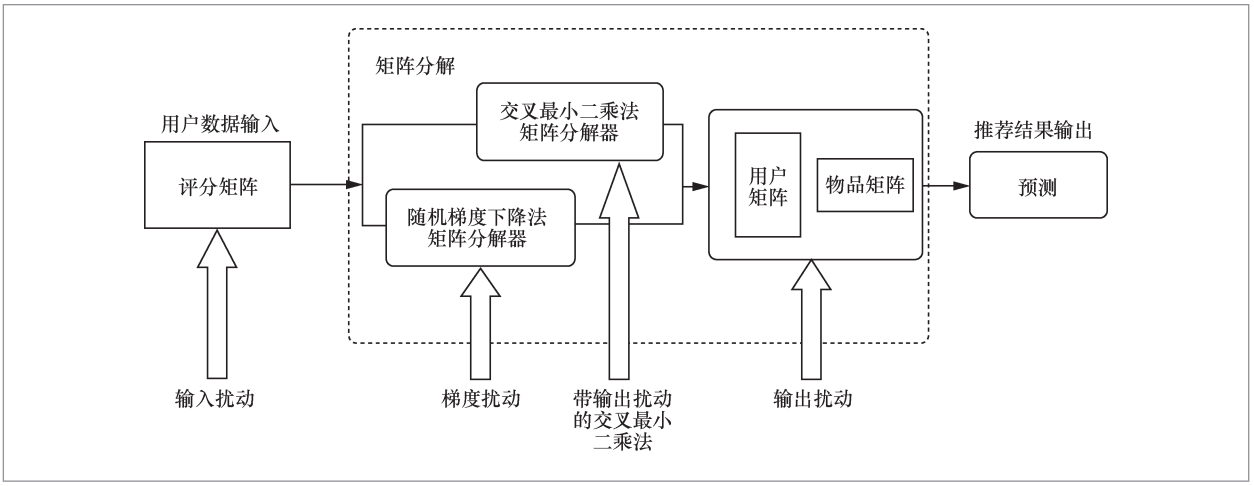


图 4 矩阵分解机制的噪声添加策略

于联邦随机梯度下降（federated stochastic gradient descent）的两阶段本地差分隐私框架FedSel，首先对本地客户端的梯度更新进行维度选择再执行数据扰动，证明了top-k梯度处理比已有的Top-1处理方法更优。Triastcyn A等人[70]使用贝叶斯计数方法代替时刻计数方法，从理论分析与实验结果两方面证明了贝叶斯差分隐私联邦学习对客户端与具体实例的隐私保护的有效性。Bonawitz K等人[71]运用秘密共享等技术，为用户向参数服务器上传参数的过程提供了隐私保证。Li T等人[72]设计了一个执行差分隐私机制的联邦推荐系统。

在推荐系统无所不在的网络环境中，用户越来越强烈地意识到自己的数据是需要保密的。通过部署更先进的隐私保护推荐系统，消除参与方对隐私泄露的担忧，对于促进数字化实体合作，防范化解重大政策风险无疑至关重要。

2.3 联邦推荐系统的跨领域推荐能力

联邦学习可以在不破坏隐私安全的前提下打破“数据孤岛”，充分利用不同用户域的信息，因此其在跨领域推荐方面也有着天然的优势。

Liu S C等人[73]通过部署不同领域的多个服务器，使用变分推理框架进行优化，最大限度地提高用户编码和所有交互域中特定领域的用户信息之间的互信息，提出了一种联邦跨领域推荐框架FedCT，在降低通信成本的同时，提出了一种不依赖于涉及域的数量、传输模型对其他领域不造成显著干扰的推理机制。此外，针对推荐系统中的冷启动问题，王健宗等人[74]基于安全内积协议设计了一种联邦协同过滤的冷启动解决方案，在数据库中添加新用户或新物品时，利用安全内积方法，联合多方评分矩阵，求解多方数据的相似矩阵，从而完成新项目的推荐输出。Wang L等人[75]则针对用户兴趣点推荐任务提出了跨领域的隐私保护联邦推荐算法，利用辅领域的用户数据对主领域用户进行兴趣分析，在用户评论分析任务上取得了较CNN更优的推荐效果，有效解决了冷启动难题。

基于跨模态的检索方法需要大量的训练数据，然而聚合、收集大量的数据将会产生巨大的隐私风险和高昂的维护成本。对此，Zong L L等人[76]提出了一种联邦跨模态训练方法，使用各模态客户端的本地数据训练各自的公共空间，再由可信服务器聚合公共子空间，并指导服务器本地模型的公共子空间更新。此项工作为结合多种数据来源形式的联邦推荐系统奠定了基础，扩宽了可使用推荐数据的范围，降低了数据收集难度。而在特定的医学应用领域，Ma J等人[77]提出了一种通信高效的联邦广义张量因子分解，以适合不同类型的实际数据，并且能够大幅度降低上行通信成本，此成果为联邦推荐在医疗领域的应用提供了基本方法。

联邦学习的引入为共同训练推荐系统的多方数据持有者提供了数据安全上的保证，以原始数据不出本地、共同利用梯度更新模型的训练方法，吸引了不同服务提供商分享不同存储形式的不同领域的用户数据，达到促进己方业务增长的目的。

2.4 联邦推荐系统的数据异质性挑战

现有的机器学习任务经常默认训练数据遵循独立同分布（independently and identically distributed，IID）假设，但是对于分布式计算而言，由于现实世界中的不同数字化实体往往服务于不同属性（如区域、爱好、财富等）的用户，其各项特征往往相差较大，存储方式也不尽相同，并且由于存在用户重合或有联系的情况，数

据相关性几乎无处不在。若是只考虑IID数据，基于现有的机器学习算法和框架训练模型，会导致许多负面效果，比如模型准确度大幅降低、无法收敛等。解决异质问题对于解决推荐系统的用户数据稀疏问题也十分重要。挖掘其他数据可以有效提高推荐准确性，用户曾经发表的文本、图像、互动信息等数据是异质的，难以处理又关乎性能的异质性数据给算法带来了挑战。

近年来，推荐系统中的非独立同分布（non-independently and identically distribution，Non-IID）情况引起了学者的重视，其不仅包含用户信息与物品信息各自非独立同分布的情况，还包括用户与物品两两在不同层级上的耦合。Cao L B[78]对推荐系统中不同的异质情况做了详细研究，并提出了判别推荐系统异质情况的理论框架。对于ICF中的物品相似度耦合，Wang C等人[79-80]首先合并了耦合的物品相似度，引入一个耦合k模式算法来预测评分，随后在矩阵分解的目标函数中添加相似性度量，以学习不易察觉的用户与项目的关系矩阵。

联邦学习在分布式学习过程中保持对设备或用户个人信息的隔离，因此异质性数据对联邦学习推理性能的影响尤其严重[81]。Wu Q等人[82]将联邦学习在实际应用中面临的异质性问题总结如下：①各个客户端在存储容量、计算能力和通信速率方面存在设备异质性问题；②各个客户端所存储数据的非独立同分布状态导致的统计异质性问题；③各个客户端所处的不同应用场景导致的模型异质性问题。数据异质性问题又可分为特征分布偏斜（协变量偏移）、标签分布偏斜（先验概率偏移）、特征与标签在不同数据来源上的不匹配、数量偏斜或不平衡等问题。笔者认为在推荐系统这一具体应用场景，可以讨论能否适当放宽联邦学习对每一轮参与设备选择的严苛要求，以解决实际遇到的异质性难题。

对于减轻异质性带来的影响，一种有效的方法是设计个性化模型，分别针对设备、数据和模型级别的异质性进行不同的个性化处理。

Yang C X等人[83]通过实证研究量化了设备异质性对联邦学习训练过程的影响。结果表明，异质性会导致联邦学习的性能下降，包括高达9.2%的准确率下降、2.32倍的训练时间延长，以及公平性遭到破坏。其中设备故障和参与者偏差是性能下降的两个潜在因素。针对设备异质性难题，Liu L M等人[84]设计了一种客户机-边缘-云分层联邦学习架构，通过引入中间边缘服务器，同时减少模型训练时间和终端设备的能耗。Xie C等人[85]则提出了一种新的异步联邦优化算法，并证明了该算法对于强凸问题以及一类受限的非凸问题具备逼近全局最优的线性收敛性。

针对统计异质性难题，McMahan B等人[44]提出了一种基于迭代平均的深层网络联邦学习方法——联邦平均（federated averaging，FedAvg），FedAvg相比联邦随机梯度下降（federated stochastic gradient descent，FedSGD）算法在通信轮次上取得了较大改善；Li X等人[86]从理论角度证明了FedAvg在处理Non-IID数据时可实现收敛，证明了此方法的有效性；Zhao Y等人[87]则进一步分析和改进了FedAvg，为了衡量客户端存储的数据分布与中央服务器掌握的数据总体分布之间的差异，引入推土机距离（earth mover's distance，EMD）来计算权重散度，表明可使用权重散度指标表示准确度下降的趋势，同时提出了一种数据共享策略，通过在初始化阶段将全局共享数据子集的随机部分分发给每个客户端，改进了 FedAvg的性能，实验证明，在不影响整体通信效率，也不增加隐私安全性漏洞的前提下，该方法可显著提高数据严重倾斜的情况的准确度。

针对模型异质性问题，Kulkarni V等人[88]总结了如下几类模型异质性个性化学习方法构建策略：增加用户上下文、迁移学习、多任务学习、元学习、知识蒸馏、基本层与个性化层共同作用、混合全局和局部模型等，并在其综述中进行了详细介绍。

此外，Wu J Z等人[89]考虑到用户本地数据通常包含公共信息（物品标签）与敏感信息（本地用户交互过的物品清单），提出了隐私异质性。相应地，为了解决隐私异质性问题，作者利用层次信息来划分公共数据和隐私数据，设计了包含公共组件与隐私组件的本地个性化模型GUM（可以实现模型的异质性），并对二者执行不同的聚合更新策略：客户端将公共组件直接发送给服务器，服务器将当前模型在本地验证集上的准确度进行加权聚合后得到新一轮的全局公共组件；而隐私组件则通过对客户端发送给服务器的本地聚类中心（视为本地用户表示的草图，不泄露本地用户数据）再进行一次聚类得到公共聚类中心，即可得到全局隐私组件，二者共同构成新一轮的全局模型。服务器更新并发送全局组件后，再由客户端以细粒度更新策略个性化更新本地模型（此方法可以解决统计异质性问题）。

2.5 联邦推荐系统的通信成本挑战

在联邦推荐系统中，特别是跨设备联邦学习设定下，联邦网络可能由大量的设备共同组成，因此每轮更新都需要进行大规模的通信，这对网络带宽与设备情况提出了较高的要求。因此，在不降低推荐准确度的前提下提升通信效率就显得尤为重要。通信开销的降低与学习效率的提升将为更大范围的工业化联邦推荐系统奠定基础。

当前设备在计算、内存和通信方面的资源有限，因此存在一些具有实用价值的不同通信优化目标。

- 压缩梯度信息。使用降维、梯度压缩等方法，减小从客户端到服务器通信对象的规模，该对象用于更新全局模型。例如Rothchild D等人[90]提出了一种使用计算草图压缩客户端更新信息大小的方法FetchSGD，将动量与累计误差从本地客户端转移到服务器上进行聚合，该方法在稀疏客户端参与时仍能保证高压缩率和良好的收敛性。Reisizadeh A等人[91]则采用周期平均和量化处理压缩客户端模型的更新信息，提出了FedPAQ（federated periodic averaging and quantization）方法，具体做法是客户端执行数次本地更新计算后，才使用量化算子向中心服务器发送此时本地结果与全局模型的差值的量化结果，并由中心服务器反量化解码后用于生成新的全局模型。该方法虽然能有效降低通信开销，但降低了收敛准确度，需要更多次训练迭代。
- 压缩广播模型。减小从服务器到客户端的全局模型广播的规模，客户端从该模型开始本地训练。例如Khan F K等人[92]通过将强化学习中的多臂老虎机算法应用于负载优化，设计奖励函数，挑选奖励反馈为正反馈的物品向量进行更新，每次只广播全局模型的一部分，在高度稀疏的推荐数据集上减少了90%的模型负载。
- 减少本地计算。修改训练算法，使得本地训练过程在计算上更加高效。例如Malinovsky G等人[93]从不动点方法的角度分析，将优化问题转化为梯度下降算子的不动点寻找问题，限制客户端的本地计算，从而突破通信开销瓶颈。

而由于联邦推荐传输的矩阵通常十分巨大，在通信成本上的考量更为关键。将参考文献[94]中提到的取子样本法或概率量化法移植到联邦推荐系统的用户矩阵上，对于大公司使用联邦推荐具有极大的吸引力。具体地，**表3**总结了联邦推荐系统对推荐社区的贡献与落地过程中遇到的挑战。

表 3 联邦推荐系统的贡献与挑战

贡献与挑战	方面	方法	参考文献
联邦推荐系统的贡献	隐私保护	加密方法	[53,71]
		数据扰动	[63,65,68-72]
	跨领域推荐	跨领域用户信息	[73,75]
		跨模态用户数据	[76-77]
落地过程中的挑战	数据异质性	设备异质性	[83-85]
		统计异质性	[1,86-87]
		模型异质性	[88-89,95-100]
	通信成本	通信优化方法	[90,94,101-104]

关于联邦优化的目标函数的研究在分散优化领域也产生了诸多成果。为了降低通信开销，Wang J Y等人[101]提出了周期性去中心化SGD方法，该方法使用多个本地更新对中心化SGD进行联邦平均。随后Li X等人[102]将该算法扩展到Non-IID数据来源中。Liang P P等人[103]提出结合本地与全局的方式，降低联邦学习通信开销，提升学习效率，该方法在参与方拥有Non-IID数据的情况下仍然有效。针对特定的应用场景，Liu Y等人[104]提出针对纵向联邦降低通信开销，进而提升学习效率的方法，该方法在理论分析与实验验证中被证明行之有效。马嘉华等人[105]则针对节点数据分布差异给联邦学习算法性能带来不良影响的问题，提出了一个基于标签量信息的节点选择算法，降低了全局模型的权重偏移上界，从而提高算法的收敛稳定性。另外，随着5G基础设备的全面铺开，联邦推荐使用的边缘设备也将处于更优的网络环境中，可以从提高速率和带宽方面正面解决通信开销问题，突破通信瓶颈。

2.6 联邦推荐系统部署

2.6.1 联邦协同过滤框架

来自华为芬兰研发中心的Ammad-ud-din M等人[106]提出了第一个基于用户隐性反馈的联邦个性化推荐系统，并展示了该方法对基准数据集MovieLens和内部数据集的适用性。作者提出了一种联邦学习框架下的隐性反馈数据集的联邦协同过滤（federated collaborative filtering，FCF）方法。该方法具有通用性，可以扩展到各种推荐系统的应用场景。FCF在中央服务器上更新主模型$\boldsymbol{Y}$（物品因子矩阵），并将其分发到每个客户端，每个用户规格模型$\boldsymbol{X}$（用户因子矩阵）仍在本地客户端，利用本地用户数据和中央服务器的$\boldsymbol{Y}$在客户端上进行更新。联邦协同过滤的基本架构如图5所示。

FCF主要考虑隐性反馈的情况，用户u对物品i偏好得分的预测可以表示为：

$$\hat{r}_{ui} = \boldsymbol{x}_u^{\top} \boldsymbol{y}_i \tag{2}$$

引入一组二元变量来表示用户u对物品i的偏好：

$$p_{ui} = \begin{cases} 1, & r_{ui} > 0 \\ 0, & r_{ui} = 0 \end{cases} \tag{3}$$

在隐性反馈的情况下，$r_{ui} = 0$可以有多种解释，比如用户u对物品i不感兴趣，或者用户u可能不知道物品i的存在等。为了解决这种不确定性问题，引入了一个确证参数：

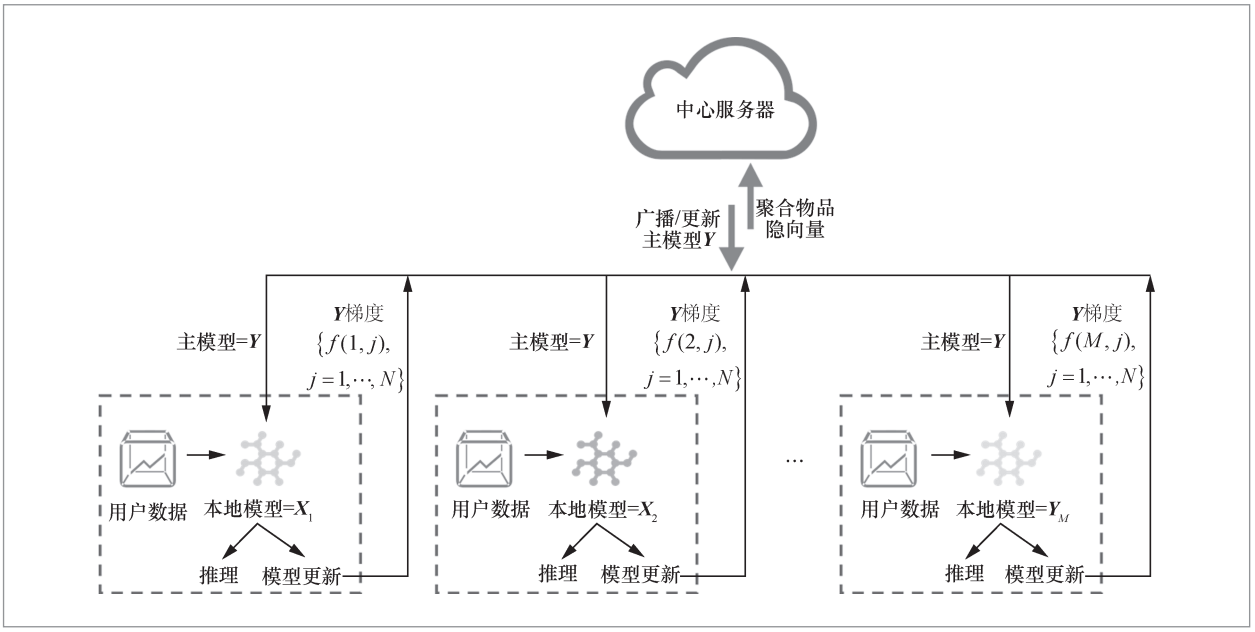


图 5 联邦协同过滤的基本架构

$$c_{ui}=1+\alpha r_{ui},\ \alpha>0 \tag{4}$$

得到所有用户u和物品i的优化目标函数：

$$J=\sum_u\sum_i c_{ui}(p_{ui}-\boldsymbol{x}_u^{\top}\boldsymbol{y}_i)^2+\lambda(\sum_u\|\boldsymbol{x}_u\|^2+\sum_i\|\boldsymbol{y}_i\|^2) \tag{5}$$

其中，λ为正则化系数。式(6)显示了对$\boldsymbol{x}_u$的最佳估计，此项可直接在用户本地计算：

$$\boldsymbol{x}_u^*=(\boldsymbol{Y}C^u\boldsymbol{Y}^{\top}+\lambda\boldsymbol{I})^{-1}\boldsymbol{Y}C^u p(u) \tag{6}$$

其中，C^u为用户u的置信参数列，$p(u)$为用户u的偏好程度。而后，为了得到物品因子$\boldsymbol{y}_i$的最佳估计值，需要知道用户因子向量$\boldsymbol{x}_i$与物品的交互信息，因此$\boldsymbol{y}_i$的更新不能在客户端完成，必须在中心服务器上完成。但从保护用户隐私的角度出发，用户与物品的交互信息只能保存在客户端设备中，因此不能直接计算$\boldsymbol{y}_i$。

为了解决这个问题，FCF提出了一种随机梯度下降的方法，在保护用户隐私的同时，允许$\boldsymbol{y}_i$在中央服务器上更新。具体来说，它使用以下计算式更新中央服务器上的$\boldsymbol{y}_i$：

$$\frac{\partial J}{\partial \boldsymbol{y}_i}=-2\sum_u[c_{ui}(p_{ui}-\boldsymbol{x}_u^{\top}\boldsymbol{y}_i)]\boldsymbol{x}_u+2\lambda\boldsymbol{y}_i \tag{7}$$

并定义：

$$f(u,i)=[c_{ui}(p_{ui}-\boldsymbol{x}_u^{\top}\boldsymbol{y}_i)]\boldsymbol{x}_u \tag{8}$$

在每个客户端u上分别计算$f(u,i)$，敏感评分信息不出本地。随后，所有客户端仅将$f(u,i)$的值发送给中心服务器进行求和，从而更新主服务器上的$\boldsymbol{y}_i$。

在通用电影推荐数据集MovieLens与其私有数据集上的实验结果表明，在不失通用性的前提下，可以认为FCF与中心化推荐方法的推荐性能表现是十分接近的。但是这一方法存在两个关键问题。一是，FCF要求每个用户和项目都参与到学习过程中来训练自己的向量，这在实际推荐场景中并不适用，因为有些用户受到设备、网络性能等限制，无法参与每一轮的

模型训练，这也与联邦学习利用用户的闲置算力的初衷相违背。二是，FCF使用物品的ID来表示物品，而对于不断上架新物品的推荐系统，无法实时扩充标记。在推荐系统的实际应用场景中，这将会带来严重的冷启动问题。

此外，在安全性方面，Chai D等人[107]指出，如果已知任意两步更新的梯度信息以及用户因子的更新计算式，可以通过求解高阶方程组推导出用户评分信息，即梯度信息有可能泄露用户隐私数据。为了解决这一问题，他们提出了一个矩阵因子分解模型——安全联邦矩阵分解（secure federated matrix factorization，SFMF）。通过对服务器与客户端之间的通信过程使用同态加密方法，杜绝了梯度泄露用户信息的可能。他们还提出了PartText传输策略：客户端只上传用户交互过的物品的梯度信息，如此在只泄露用户交互物品列表的情况下就能获得计算时间上的较大改善。Minto L等人[108]则对隐反馈FCF加以改进，通过在客户端上传的物品更新梯度矩阵中加入本地差分隐私策略以及代理网络来得到不包含用户元数据的物品更新梯度矩阵，以获得更强的隐私保护力度，防止第三方获取物品更新梯度矩阵后实施重构攻击。

隐反馈数据在进行联邦化改造时具有天然优势：由于将所有与用户无交集的物品都当成负样本进行求导，服务器无法探知用户的交集物品列表。而FCF对显反馈数据的求导式子中只包含该用户评分过的物品，由此容易泄露用户交互记录。对于此弊端，Lin G Y等人[109]提出将FedRec扩展到了显性反馈的推荐场景，在上传用户梯度时将随机采样的部分未评分物品一起上传到服务器以遮掩用户实际交互的物品信息，并且采用用户平均评分与混合评分机制来生成负采样物品的评分。但此举引入了额外的噪声，因此Liang F等人[110]又进一步提出FedRec++，分配部分去噪客户端以隐私感知的方式消除噪声数据，获得了更优的推荐性能。

2.6.2 应用于新闻推荐任务的FCF

针对FCF无法处理新项目和具有梯度泄露风险的问题，Qi T等人[111]对新闻推荐任务的FCF应用进行了改进，提出FedNewsRec方法，在去中心化存储条件下，基于用户行为数据训练了一个精确的新闻推荐模型。具体地，Qi T等人[111]采用从低到高由词嵌入层、卷积神经网络、多头自注意力网络、注意力网络构成的4层新闻模型学习新闻表征，并使用参考文献[112]中的用户模型学习用户的点击历史，如**图6**所示。

在FedNewsRec框架中，每一轮更新时，一组随机选择的用户的局部梯度被上传到服务器，然后进一步聚合以更新服务器中的全局模型。提供新闻服务的服务器不记录或收集用户行为，这可以解决隐私问题，降低数据泄露的风险。此外，在客户端和中心服务器之间的通信过程中，额外引入本地差分隐私技术来保护上传梯度中的隐私信息（但这降低了聚合梯度模型的更新精度）。Qi T等人[111]将FedNewsRec与当前主流的几种新闻推荐方法在新闻数据集Adressa和MSN-News上进行了比较，验证了FedNewsRec在个性化新闻推荐模型学习中的性能，其优于FCF而稍逊于中心新闻推荐模型。

相比于FCF，FedNewsRec可以处理新用户和新项目，并且不需要所有用户参与到训练中，但由于引入了新闻推荐模型，FedNewsRec 不适用于其他场景，不具备普适性。

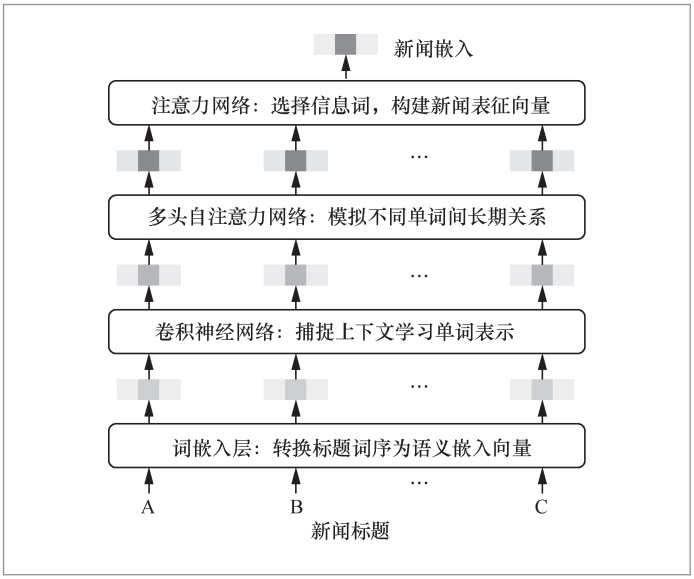


图 6 FedNewsRec 的新闻表征向量模型

2.6.3 联邦多视图推荐框架

针对FedNewsRec的适用性缺点，Huang M K等人[113]结合深度结构化语义模型(deep structured semantic model, DSSM)[114]与FCF，进一步提出了一个基于内容的通用联邦多视图推荐框架FL-MV-DSSM(federated learning-multi view-deep structured semantic model)，该框架通过使用应用的信息训练一个共享的用户子模型，从而实现更好的物品推荐性能。

FL-MV-DSSM可以处理现有的FedRec冷启动问题，通过将一般的深度结构化语义模型转变成联邦学习环境，FL-MV-DSSM可以将用户和物品映射到一个共享的语义空间中，进一步实现基于内容的推荐。在此基础上，从多个数据源学习联邦模型，以获得更丰富的用户级特征，提高了FL-MV-DSSM的推荐性能。此外，FL-MV-DSSM还提供了一种新的联邦多视图设置，所有视图都协同训练一个模型，且视图之间不存在原始数据交互，每个视图对物品子模型的贡献也受到保护，通过加密或对可信执行环境进行视图级别隔离，恶意视图无法通过监视其对共享局部物品子模型的更改，从梯度中推断出正常视图的原始数据。

2.6.4 基于元学习的联邦个性化推荐框架

FedNewsRec获取更丰富用户个性化信息的思想与元学习的思想不谋而合，相对应地，Lin Y J等人[115]提出了一种联邦元学习推荐系统框架元矩阵分解MetaMF，针对不同的客户端，学习一个较小规模的个性化本地模型，在减少资源消耗的同时，获得高于基线方法的预测准确度，且对新用户的适应只需要几个更新步骤。此外，为了得到更优的推荐性能，还可以像参考文献[34]那样使用更强大的深度因子分解机模型。

2.7 联邦推荐系统的优化

在原来的联邦推荐算法中，客户端的选择是随机的，更新梯度信息的聚合也只是简单的加权平均过程(称为FedAvg[48])。在提高收敛速度和减少带宽消耗上还有很大的改进空间。

Muhammad K等人[116]提出了一种联邦学习算法FedFast，以提升FedAvg的效率，提高联邦推荐系统的表现。其核心创新点是提出了两种算法：①ActvSAMP可以基于聚类算法选择更具代表性的训练参与者；②ActvAGG对不同嵌入采取不同的聚合方式，合并训练模型，加速模型收敛。

ActvSAMP算法将K个参与者划分为p组，每轮选取m个参与者参加本轮训练。参考文献[111]使用其他具有隐私保护属

性的特征（如所处地区、设备类型等）进行聚类，将具有相似嵌入的参与者划分到同一个聚类中。随后从每个聚类中挑选出区别较大的客户端作为代表用户参与训练。

有别于常规的FedAvg仅对局部模型参数进行加权平均来更新全局模型的权重矩阵的做法，FedFast使用ActvAGG算法对3个部分的参数进行更新：用户嵌入、物品嵌入、模型权重矩阵。在用户嵌入的学习过程中，任何代表用户学习到的参数更新，都将在乘以大小随迭代轮次 t 降低的折扣系数 $\exp(-t)$ 后应用到与其处于相同聚类的从属用户上。这样就可以在仅计算少量用户的梯度更新的前提下，有效加快最终收敛，减少单轮算力消耗。

但鉴于参考文献[116]的基线算法只有FedAvg，比较意义有限。此外，该方法使用的聚类算法时间复杂度较高，影响了模型全局表现，且在通信轮次和每轮通信流量之间进行了权衡，这意味着更新时需要额外的网络开销。具体地，**表4**给出了联邦推荐系统部署实践的优劣势对比。

2.8 其他联邦推荐系统实际应用

针对推荐系统的另一个子课题排序学习（learning to rank，LTR），Kharitonov E[117]设计了一个基于差分隐私的隐私保护联邦在线学习排名（federated online learning to rank，FOLtR）系统，用于对在线学习的效果进行评估，取得了接近基线方法的水准，且具备一定的处理噪声隐私信息的能力。

Trienes J等人[118]提出在去中心化的社交网络上使用推荐算法，以解决大规模用户监视和滥用用户数据影响选举公平的问题。作者使用联邦环境社交网络收集大量无偏样本，分别结合协同过滤和拓扑图设计了相应的推荐器，并证实协同过滤方法优于拓扑方法。

更具实操性地，Tan B等人[119]部署了一个实用的联邦推荐框架，实现了大量流行算法，支持各种在线推荐服务。该系统由数据层、算法层、服务层和接口层组成，支

表4 联邦推荐系统部署实践的优劣势

模型	介绍	解决问题	局限性
FCF[106]	针对隐反馈数据首先提出联邦协同过滤算法	在保护隐私数据的前提下共同训练协同过滤推荐模型	需要所有客户端参与；存在冷启动问题；存在梯度泄露风险
SFMF[107]	加入同态加密进行安全矩阵分解	防止梯度泄露	在隐私保护力度与算力开销之间难以两全
Stronger FCF[108]	使用本地差分隐私与代理服务器保护物品更新梯度矩阵	防止第三方获取物品更新梯度矩阵实施重构攻击	仅在隐反馈推荐情景下证明了有效性
FedRec[109]、FedRec++[110]	适用于显反馈评分数据的联邦协同过滤，后者增加了消除噪声的机制	对基于显反馈数据的推荐系统进行了联邦化实践	与现代推荐算法的结合不多
FedNewsRec[111]	结合深度网络学习新闻表征与用户历史，产生新闻推荐	解决冷启动问题；不需要所有用户参与	非通用模型
FL-MV-DSSM[113]	多视图深度结构化语义模型	解决冷启动问题；多视图	带来新的安全挑战
MetaMF[115]	引入元学习矩阵分解训练个性化模型	快速适应新用户	仅适用于评分预测任务
FedFast[116]	在客户端选择与更新信息聚合方面提出改善方法	加快模型收敛，减少计算量	带来额外的通信开销

持内容推荐、产品推荐、在线广告等各种在线应用。算法层有通用矩阵因子化、SVD、因子化机、广度与深度学习等方法。在联邦学习的环境下，将算法层建立在FATE框架上，并发布在线内容推荐演示。

3 基于联邦学习的推荐系统研究方向展望

本文讨论了目前联邦学习与推荐系统的结合情况，并对基于联邦学习的实现方法进行了大致的分类梳理，在文献调研过程中，发现目前二者的结合工作还不是很紧密，主要集中于对传统的推荐模型（如协同过滤）进行联邦化改造上，联邦学习在推荐系统中的应用仍然处于起步阶段，实际应用层面需要更多的研究与讨论。

下面笔者对推荐系统场景下的联邦学习与传统技术的结合提出6个可能的研究方向。

（1）缺少部分用户数据时的可解释性问题

推荐系统的可解释性指在给予用户推荐结果的同时，展示对结果的支持论据或推荐解释，以降低用户反感度，增强推荐说服力。由于使用了其他参与方的数据共同生成推荐结果，传统的提高推荐系统解释性的方法不再适用，如何在缺少他方原始数据的情况下生成可信度高的解释结果，降低用户反感度，依然是比较冷门的研究方向。由于联邦推荐系统严格保护本地用户信息，无法直接基于用户特征生成可读性解释，可以考虑利用知识图谱等技术建立起本地用户特征与全局物品特征之间的关联关系，对用户进行推荐与解释。此外，由于联邦推荐系统在不同的客户端上可以采用个性化推荐模型，亟须设计与模型无关的可解释推荐框架，在不传输客户端敏感信息、不逐一设计解释方案的前提下，给出全局推荐说明，以提高传统的推荐系统可解释性增强方法的可扩展性。

（2）联邦推荐系统的安全性证明与维护设想

传统的安全攻击与防御方法可以应用于联邦推荐场景的不同进程当中，如Ribero M等人[120]详细阐述了差分隐私算法在联邦推荐系统通信环节中的应用，Hu H S等人[121]则在本地处理环节引入了局部敏感哈希。针对恶意参与方污染或攻击聚合环节，Blanchard P等人[122]提出的Krum方法与Mhamdi E M E等人[123]提出的Bulyan模型参数聚合方法可以有效防御拜占庭攻击，但均牺牲了收敛速度和准确率，对每轮参与者的数量以及计算复杂度也有更高要求。具体地，针对联邦推荐场景研究更有效的防御方法，对于增强参与方的互信具有重大现实意义，值得更多的研究者关注。

（3）对更多深度学习推荐系统的联邦化改造

深度学习技术在推荐系统上已经得到了十分广泛的应用，采用深层神经网络结构的推荐模型能够学习到更抽象、更稠密的用户与项目的表示，以及二者交互的非线性结构特征，产生更精准的推荐结果。但是具体到与联邦学习结合的场景，大部分工作仍停留于协同过滤等传统推荐算法，对于各种深度模型的结合效果，特别是不同神经网络的梯度传输与聚合，需要不同的标准重新进行评估，以及继续探索联邦深度推荐系统的可能性。Wang H等人[124]针对各种现代神经网络（如卷积神经网络、长短期记忆网络）的联邦学习环境，提出了通过匹配和平均隐藏元素，以分层方式构建共享全局模型的训练方法

FedMA。这一方法有望被应用于更多使用复杂神经网络的推荐系统的联邦化改造上。

（4）在保护隐私的前提下充分利用用户数据

现有的联邦推荐系统大多仅使用了用户的物品交互信息，并在“数据孤岛”模拟方式上采取了简单的比例分割法，而现实情况往往更加复杂，诸如需要多少数据量才能共同训练出较精准的推荐系统、数据在客户端间的不同分布情况会如何影响联邦推荐系统的性能、传统推荐器使用的额外数据（社交数据、时空数据等）在联邦推荐系统中是否仍然有效等问题仍未得到解答。

（5）制定符合实际的参与方贡献评估策略

由于联邦推荐系统使用多方数据共同产生推荐结果，不同客户端从这一联合训练过程中得到的收益并不相等，因此需要制定公平的定价策略，科学评估参与方的贡献，协调相关方的利益。传统的特征重要性评估方法只能评估全局特征的重要性，在联邦推荐系统上无法直接使用。最初的联邦推荐系统贡献评估方法[125-126]多使用博弈论中的沙普利值（Shapley value，SV）作为衡量指标，但其计算复杂度过高，严重影响了模型收敛速度，增加了通信成本，并且其数值与使用的推荐模型强相关，同一参与用户在不同的模型下的SV也会不同[127]。对于联邦推荐场景，至今仍缺少多数人认同的贡献评估策略，这对于联邦推荐系统的应用落地仍是巨大的阻碍，是值得学术界与工业界关注的研究方向。

（6）利用联邦学习进行跨领域信息融合的推荐

随着数据时代的到来，用户无时无刻不在多个领域产生数据，但现实中融合同一用户在不同平台上的数据进行跨领域推荐往往止步于平台间的不互信甚至对立，用户隐私保护法规往往也限制了多领域知识的融合。借助联邦学习打破“数据孤岛”的能力，并结合深度学习技术，可以将不同平台间的各类数据以嵌入式表示等方法构建深层预测模型，并共同训练产生令人满意的推荐结果。未来，如何更好地利用联邦学习融合跨领域信息，有效缓解数据稀疏问题与冷启动问题，将是联邦推荐系统的重点研究方向。

4 结束语

现有文献在发展联邦推荐系统方面已经付出了很多努力，对当前的联邦推荐系统进行完整的概述和总结是很有意义的。受以前的联邦系统的启发，笔者概述了基于联邦学习的推荐，讨论了几种最新的联邦推荐系统的独特属性和相关的挑战。在此基础上，还对现有的联邦推荐系统的特点和设计进行了比较，并对联邦推荐场景的研究方向做了展望。可以预见，在不久的将来，联邦推荐系统将打破壁垒，充分利用所有的数据，在保护各方数据安全的前提下进行准确的推荐，人们获取目标信息的效率将发生巨大的变化。

参考文献:

[1] ZHANG S, YAO L N, SUN A X, et al. Deep learning based recommender system[J]. ACM Computing Surveys, 2020, 52(1): 1-38.

[2] RICCI F, ROKACH L, SHAPIRA B. Introduction to recommender systems handbook[M]//Recommender Systems Handbook. Heidelberg: Springer, 2011.

[3] ALBRECHT J P. How the GDPR will change the world[J]. European Data Protection Law Review, 2016, 2(3): 287–289.

[4] MOONEY R J, ROY L. Content–based book recommending using learning for text categorization[C]//Proceedings of the 5th ACM Conference on Digital Libraries. New York: ACM Press, 2000: 195–204.

[5] 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. New York: ACM Press, 1995: 210–217.

[6] MILLER B N, ALBERT I, LAM S K, et al. MovieLens unplugged: experiences with an occasionally connected recommender system[C]//Proceedings of the 8th International Conference on Intelligent User Interfaces. New York: ACM Press, 2003: 263–266.

[7] TERVEEN L, HILL W, AMENTO B, et al. Phoaks: a system for sharing recommendations[J]. Communications of the ACM, 1997, 40(3): 59–62.

[8] GOLDBERG K, ROEDER T, GUPTA D, et al. Eigentaste: a constant time collaborative filtering algorithm[J]. Information Retrieval, 2001, 4(2): 133–151.

[9] BILLSUS D, PAZZANI M J. Learning collaborative information filters[C]// Proceedings of the 15th International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc., 1998: 46–54.

[10] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: a survey of the state–of–the–art and possible extensions[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6): 734–749.

[11] DESHPANDE M, KARYPIS G. Item–based top–N recommendation algorithms[J]. ACM Transactions on Information Systems, 2004, 22(1): 143–177.

[12] SU X Y, KHOSHGOFTAAR T M. A survey of collaborative filtering techniques[J]. Advances in Artificial Intelligence, 2009: 421425.

[13] CHOI K, YOO D, KIM G, et al. A hybrid online–product recommendation system: combining implicit rating–based collaborative filtering and sequential pattern analysis[J]. Electronic Commerce Research and Applications, 2012, 11(4): 309–317.

[14] XIANG L, YUAN Q, ZHAO S W, et al. Temporal recommendation on graphs via long– and short–term preference fusion[C]//Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2010: 723–732.

[15] BREESE J S, HECKERMAN D, KADIE C. Empirical analysis of predictive algorithms for collaborative filtering[C]//Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers Inc., 1998: 43–52.

[16] UNGAR L H, FOSTER D P. Clustering methods for collaborative filtering[C]// Proceedings of 1998 AAAI Workshop on Recommendation Systems. Palo Alto: AAAI Press, 1998: 114–129.

[17] 黄立威, 江碧涛, 吕守业, 等. 基于深度学习的推荐系统研究综述[J]. 计算机学报, 2018, 41(7): 1619–1647.
HUANG L W, JIANG B T, LYU S Y, et al. Survey on deep learning based recommender systems[J]. Chinese Journal of Computers, 2018, 41(7): 1619–1647.

[18] PAZZANI M J, BILLSUS D. Content–based recommendation systems[M]//The adaptive Web. Heidelberg: Springer, 2007.

[19] SALTON G, WONG A, YANG C S. A vector space model for automatic indexing[J]. Communications of the ACM,

1975, 18(11): 613–620.

[20] BURKE R. Hybrid recommender systems: survey and experiments[J]. User Modeling and User–Adapted Interaction, 2002, 12(4): 331–370.

[21] WANG H, WANG N Y, YEUNG D Y. Collaborative deep learning for recommender systems[C]//Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2015: 1235–1244.

[22] YANG X W, STECK H, LIU Y. Circle–based recommendation in online social networks[C]//Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2012: 1267–1275.

[23] KRULWICH B. LIFESTYLE FINDER: intelligent user profiling using large–scale demographic data[J]. AI Magazine, 1997, 18(2): 37–45.

[24] GONZÁLEZ G, LOPEZ B, ROSA J L. A multi–agent smart user model for cross–domain recommender systems[C]// Proceedings of 2005 International Conference on Intelligent User Interfaces: Beyond Personalization. New York: ACM Press, 2005.

[25] DENG L. Deep learning: methods and applications[J]. Foundations and Trends in Signal Processing, 2014, 7(3/4): 197–387.

[26] 王健宗, 黄章成, 肖京. 人工智能赋能金融科技[J]. 大数据, 2018, 4(3): 111–116.
WANG J Z, HUANG Z C, XIAO J. Artificial intelligence energize Fintech[J]. Big Data Research, 2018, 4(3): 111–116.

[27] SALAKHUTDINOV R, MNIH A, HINTON G. Restricted Boltzmann machines for collaborative filtering[C]//Proceedings of the 24th International Conference on Machine Learning. New York: ACM Press, 2007: 791–798.

[28] SEDHAIN S, MENON A K, SANNER S, et al. AutoRec: autoencoders meet collaborative filtering[C]//Proceedings of the 24th International Conference on World Wide Web. New York: ACM Press, 2015: 111–112.

[29] HE X N, LIAO L Z, ZHANG H W, et al. Neural collaborative filtering[C]//Proceedings of the 26th International Conference on World Wide Web. New York: ACM Press, 2017: 173–182.

[30] ZHOU G R, ZHU X Q, SONG C R, et al. Deep interest network for click–through rate prediction[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2018: 1059–1068.

[31] ZHOU G R, MOU N, FAN Y, et al. Deep interest evolution network for click–through rate prediction[C]//Proceedings of 2019 AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2019: 5941–5948.

[32] LI C, LIU Z Y, WU M M, et al. Multi–interest network with dynamic routing for recommendation at Tmall[C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York: ACM Press, 2019: 2615–2623.

[33] HE X N, CHUA T S. Neural factorization machines for sparse predictive analytics[C]//Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2017: 355–364.

[34] GUO H F, TANG R M, YE Y M, et al. DeepFM: a factorization–machine based neural network for CTR prediction[C]// Proceedings of the 26th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2017: 1725–1731.

[35] XIAO J, YE H, HE X N, et al. Attentional factorization machines: learning the

weight of feature interactions via attention networks[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2017: 3119–3125.

[36] GONG Y Y, ZHANG Q. Hashtag recommendation using attention-based convolutional neural network[C]//Proceedings of the 25th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2016: 2782–2788.

[37] ZHANG Q, WANG J W, HUANG H R, et al. Hashtag recommendation for multimodal microblog using co-attention network[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2017: 3420–3426.

[38] OORD A V D, DIELEMAN S, SCHRAUWEN B. Deep content-based music recommendation[C]//Proceedings of the 26th International Conference on Nerual Information Processing Systems. Red Hook: Curran Associates Inc., 2013: 2643–2651.

[39] LEI C Y, LIU D, LI W P, et al. Comparative deep learning of hybrid representations for image recommendations[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 2545–2553.

[40] HIDASI B A Z, KARATZOGLOU A, BALTRUNAS L, et al. Session-based recommendations with recurrent neural networks[C]//Proceedings of 2015 International Conference on Learning Representations. [S.l.:s.n.], 2015.

[41] LIU Q, WU S, WANG L. Multi-behavioral sequential prediction with recurrent log-bilinear model[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(6): 1254–1267.

[42] WU C H, WANG J W, LIU J T, et al. Recurrent neural network based recommendation for time heterogeneous feedback[J]. Knowledge-Based Systems, 2016, 109: 90–103.

[43] LI Y, LIU T, JIANG J, et al. Hashtag recommendation with topical attention-based LSTM[C]//Proceedings of the 26th International Conference on Computational Linguistics. Cambridge: The MIT Press, 2016: 943–952.

[44] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. [S.l.:s.n.], 2017: 1273–1282.

[45] KONEČNÝ J, MCMAHAN B, RAMAGE D. Federated optimization: Distributed optimization beyond the datacenter[J]. arXiv preprint, 2015, arXiv:151103575.

[46] KAIROUZ E B P, MCMAHAN H B. Advances and open problems in federated learning[J]. Foundations and Trends in Machine Learning, 2021, 14(1).

[47] LI Q B, WEN Z Y, WU Z M, et al. Federated learning systems: vision, hype and reality for data privacy and protection[J]. arXiv preprint, 2019, arXiv:190709693.

[48] HARD A, RAO K, MATHEWS R, et al. Federated learning for mobile keyboard prediction[J]. arXiv preprint, 2018, arXiv:181103604.

[49] YANG Q, LIU Y, CHEN T J, et al. Federated machine learning[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2): 1–19.

[50] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345–1359.

[51] 吴建汉, 司世景, 王健宗, 等. 联邦学习攻击与防御综述[J]. 大数据, 2022: 2022038.
WU J H, SI S J, WANG J Z, et al. Threats

and defenses of federated learning: a survey[J]. Big Data Research, 2022: 2022038.

[52] 周俊, 董晓蕾, 曹珍富. 推荐系统的隐私保护研究进展[J]. 计算机研究与发展, 2019, 56(10): 2033-2048.
ZHOU J, DONG X L, CAO Z F. Research advances on privacy preserving in recommender systems[J]. Journal of Computer Research and Development, 2019, 56(10): 2033-2048.

[53] KIM S, KIM J, KOO D, et al. Efficient privacy-preserving matrix factorization via fully homomorphic encryption: extended abstract[C]//Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. New York: ACM Press, 2016: 617-628.

[54] AGRAWAL R, SRIKANT R. Privacy-preserving data mining[C]//Proceedings of 2000 ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2000: 439-50.

[55] POLAT H, DU W L. Privacy-preserving collaborative filtering using randomized perturbation techniques[C]//Proceedings of 3rd IEEE International Conference on Data Mining. Piscataway: IEEE Press, 2003: 625-628.

[56] HERLOCKER J, KONSTAN J, BORCHERS A, et al. An algorithmic framework for performing collaborative filtering [C]//Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 1999: 230-237.

[57] CHEN K K, LIU L. Privacy preserving data classification with rotation perturbation[C]//Proceedings of 5th IEEE International Conference on Data Mining. Piscataway: IEEE Press, 2005: 589-592.

[58] DWORK C. A firm foundation for private data analysis[J]. Communications of the ACM, 2011, 54(1): 86-95.

[59] DWORK C, ROTH A. The algorithmic foundations of differential privacy[J]. Foundations and Trends in Theoretical Computer Science, 2013, 9(3/4): 211-407.

[60] MCSHERRY F, MIRONOV I. Differentially private recommender systems: building privacy into the Netflix prize contenders[C]//Proceedings of the 15th ACM SIGKDD International Conference on Knowledge discovery and data mining. New York: ACM Press, 2009: 627-636.

[61] ZHU T Q, REN Y L, ZHOU W L, et al. An effective privacy preserving algorithm for neighborhood-based collaborative filtering[J]. Future Generation Computer Systems, 2014, 36: 142-155.

[62] 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. New York: ACM Press, 2015: 107-114.

[63] OREKONDY T, OH S J, ZHANG Y, et al. Gradient-leaks: understanding and controlling deanonymization in federated learning[J]. arXiv preprint, 2018, arXiv:180505838.

[64] WANG Z B, SONG M K, ZHANG Z F, et al. Beyond inferring class representatives: user-level privacy leakage from federated learning[C]//Proceedings of 2019 IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2019: 2512-2520.

[65] MELIS L, SONG C Z, DE CRISTOFARO E, et al. Exploiting unintended feature leakage in collaborative learning[C]//Proceedings of 2019 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2019: 691-706.

[66] 王健宗, 孔令炜, 黄章成, 等. 联邦学习隐私保护研究进展[J]. 大数据, 2021, 7(3): 130-149.
WANG J Z, KONG L W, HUANG Z C, et al. Research advances on privacy protection of federated learning[J]. Big

Data Research, 2021, 7(3): 130-149.

[67] 周传鑫, 孙奕, 汪德刚, 等. 联邦学习研究综述[J]. 网络与信息安全学报, 2021, 7(5): 77-92.
ZHOU C X, SUN Y, WANG D G, et al. Survey of federated learning research[J]. Chinese Journal of Network and Information Security, 2021, 7(5): 77-92.

[68] TRUEX S, LIU L, CHOW K H, et al. LDP-Fed: federated learning with local differential privacy[C]//Proceedings of the 3rd ACM International Workshop on Edge Systems, Analytics and Networking. New York: ACM Press, 2020: 61-66.

[69] LIU R X, CAO Y, YOSHIKAWA M, et al. FedSel: federated SGD under local differential privacy with top-k dimension selection[C]//Proceedings of the International Conference on Database Systems for Advanced Applications. Heidelberg: Springer, 2020: 485-501.

[70] TRIASTCYN A, FALTINGS B. Federated learning with Bayesian differential privacy[C]//Proceedings of 2019 IEEE International Conference on Big Data. Piscataway: IEEE Press, 2019: 2587-2596.

[71] BONAWITZ K, IVANOV V, KREUTER B, et al. Practical secure aggregation for privacy-preserving machine learning[C]//Proceedings of 2017 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2017: 1175-1191.

[72] LI T, SONG L Q, FRAGOULI C. Federated recommendation system via differential privacy[C]//Proceedings of 2020 IEEE International Symposium on Information Theory. Piscataway: IEEE Press, 2020: 2592-2597.

[73] LIU S C, XU S Y, YU W H, et al. FedCT: federated collaborative transfer for recommendation[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2021: 716-725.

[74] 王健宗, 肖京, 朱星华, 等. 联邦推荐系统的协同过滤冷启动解决方法[J]. 智能系统学报, 2021, 16(1): 178-185.
WANG J Z, XIAO J, ZHU X H, et al. Cold starts in collaborative filtering for federated recommender systems[J]. CAAI Transactions on Intelligent Systems, 2021, 16(1): 178-185.

[75] WANG L, WANG Y H, BAI Y, et al. POI recommendation with federated learning and privacy preserving in cross domain recommendation[C]//Proceedings of 2021 IEEE Conference on Computer Communications Workshops. Piscataway: IEEE Press, 2021: 1-6.

[76] ZONG L L, XIE Q J, ZHOU J H, et al. FedCMR: federated cross-modal retrieval[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2021: 1672-1676.

[77] MA J, ZHANG Q C, LOU J, et al. Communication efficient federated generalized tensor factorization for collaborative health data analytics[C]//Proceedings of the International World-Wide Web Conference. New York: ACM Press, 2021: 171-182.

[78] CAO L B. Non-IID recommender systems: a review and framework of recommendation paradigm shifting[J]. Engineering, 2016, 2(2): 212-224.

[79] WANG C, CAO L B, WANG M C, et al. Coupled nominal similarity in unsupervised learning[C]//Proceedings of the 20th ACM international conference on Information and knowledge management. New York: ACM Press, 2011: 973-978.

[80] WANG C, DONG X J, ZHOU F, et al. Coupled attribute similarity learning on categorical data[J]. IEEE Transactions on Neural Networks and Learning Systems,

2015, 26(4): 781–797.

[81] 王健宗, 孔令炜, 黄章成, 等. 联邦学习算法综述[J]. 大数据, 2020, 6(6): 64–82.
WANG J Z, KONG L W, HUANG Z C, et al. Research review of federated learning algorithms[J]. Big Data Research, 2020, 6(6): 64–82.

[82] WU Q, HE K, CHEN X. Personalized federated learning for intelligent IoT applications: a cloud–edge based framework[J]. IEEE Computer Graphics and Applications, 2020, 1: 35–44.

[83] YANG C X, WANG Q P, XU M W, et al. Characterizing impacts of heterogeneity in federated learning upon large–scale smartphone data[C]//Proceedings of 2021 International World–Wide Web Conference.. New York: ACM Press, 2021: 935–946.

[84] LIU L M, ZHANG J, SONG S H, et al. Client–edge–cloud hierarchical federated learning[C]//Proceedings of 2020 IEEE International Conference on Communications. Piscataway: IEEE Press, 2020: 1–6.

[85] XIE C, KOYEJO S, GUPTA I. Asynchronous federated optimization[J]. arXiv preprint, 2019, arXiv:1903.03934.

[86] LI X, HUANG K X, YANG W H, et al. On the convergence of FedAvg on Non–IID data[C]//Proceedings of the 7th International Conference on Learning Representations. [S.l.:s.n.],2019.

[87] ZHAO Y, LI M, LAI L Z, et al. Federated learning with Non–IID data[J]. arXiv preprint, 2018, arXiv:180600582.

[88] KULKARNI V, KULKARNI M, PANT A. Survey of personalization techniques for federated learning[C]//Proceedings of 2020 4th World Conference on Smart Trends in Systems, Security and Sustainability. Piscataway: IEEE Press, 2020: 794–797.

[89] WU J Z, LIU Q, HUANG Z Y, et al. Hierarchical personalized federated learning for user modeling[C]// Proceedings of 2021 Web Conference. New York: ACM Press, 2021: 957–968.

[90] ROTHCHILD D, PANDA A, ULLAH E, et al. FetchSGD: communication–efficient federated learning with sketching[C]// Proceedings of the 37th International Conference on Machine Learning. New York: ACM Press, 2020.

[91] REISIZADEH A, MOKHTARI A, HASSANI H, et al. FedPAQ: a communication–efficient federated learning method with periodic averaging and quantization[C]//Proceedings of 2020 International Conference on Artificial Intelligence and Statistics. [S.l.:s.n.], 2020: 2021–2031.

[92] KHAN F K, FLANAGAN A, TAN K E, et al. A payload optimization method for federated recommender systems[C]// Proceedings of the 15th ACM Conference on Recommender Systems. New York: ACM Press, 2021: 432–442.

[93] MALINOVSKY G, KOVALEV D, GASANOV E, et al. From local SGD to local fixed point methods for federated learning[C]//Proceedings of the 37th International Conference on Machine Learning. New York: ACM Press, 2020: 6692–6701.

[94] KONEČNÝ J, MCMAHAN H B, YU F X, et al. Federated learning: strategies for improving communication efficiency[C]// Proceedings of the NIPS Workshop on Private Multi–Party Machine Learning. Cambridge: The MIT Press, 2016.

[95] MANSOUR Y, MOHRI M, RO J, et al. Three approaches for personalization with applications to federated learning[J]. arXiv preprint, 2020, arXiv:2002.10619.

[96] SMITH V, CHIANG C K, SANJABI M, et al. Federated multi–task learning[C]// Proceedings of the 31st International Conference on Neural Information

Processing Systems. Red Hook: Curran Associates Inc., 2017: 4427–4437.

[97] LINDEN G, SMITH B, YORK J. Amazon.com recommendations: item-to-item collaborative filtering[J]. IEEE Internet Computing, 2003, 7(1): 76–80.

[98] LI D L, WANG J P. FedMD: heterogenous federated learning via model distillation[J]. arXiv preprint, 2019, arXiv:1910.03581.

[99] ARIVAZHAGAN M G, AGGARWAL V, SINGH A K, et al. Federated learning with personalization layers[J]. arXiv preprint, 2019, arXiv:1912.00818.

[100]HANZELY F, RICHTÁRIK P. Federated learning of a mixture of global and local models[J]. arXiv preprint, 2020, arXiv:2002.05516.

[101]WANG J Y, JOSHI G. Cooperative SGD: a unified framework for the design and analysis of communication-efficient SGD algorithms[C]//Proceedings of the ICML Workshop on Coding Theory for Machine Learning. [S.l.:s.n.],2019.

[102]LI X, YANG W H, WANG S S, et al. Communication efficient decentralized training with multiple local updates[J]. arXiv preprint, 2019, arXiv:1910.09126.

[103]LIANG P P, LIU T, ZIYIN L, et al. Think locally, act globally: federated learning with local and global representations[J]. arXiv preprint, 2020, arXiv:200101523.

[104]LIU Y, KANG Y, ZHANG X W, et al. A communication efficient vertical federated learning framework[J]. arXiv preprint, 2019, arXiv:1912.11187.

[105]马嘉华, 孙兴华, 夏文超, 等. 基于标签量信息的联邦学习节点选择算法[J]. 物联网学报, 2021, 5(4): 46–53.
MA J H, SUN X H, XIA W C, et al. Node selection based on label quantity information in federated learning[J]. Chinese Journal on Internet of Things, 2021, 5(4): 46–53.

[106]AMMAD-UD-DIN M, IVANNIKOVA E, KHAN S A, et al. Federated collaborative filtering for privacy-preserving personalized recommendation system[J]. arXiv preprint, 2019, arXiv:190109888.

[107]CHAI D, WANG L Y, CHEN K, et al. Secure federated matrix factorization[J]. IEEE Intelligent Systems, 2021, 36(5): 11–20.

[108]MINTO L, HALLER M, LIVSHITS B, et al. Stronger privacy for federated collaborative filtering with implicit feedback[C]//Proceedings of the 15th ACM Conference on Recommender Systems. New York: ACM Press, 2021: 342–350.

[109]LIN G Y, LIANG F, PAN W K, et al. FedRec: federated recommendation with explicit feedback[J]. IEEE Intelligent Systems, 2021, 36(5): 21–30.

[110]LIANG F, PAN W, MING Z. FedRec++: lossless federated recommendation with explicit feedback[C]//Proceedings of the 35th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021: 4224–4231.

[111]QI T, WU F Z, WU C H, et al. Privacy-preserving news recommendation model learning[C]//Proceedings of 2020 Conference on Empirical Methods in Natural Language Processing: Findings. [S.l.]: ACL Press, 2020: 1423–32.

[112]OKURA S, TAGAMI Y, ONO S, et al. Embedding-based news recommendation for millions of users[C]//Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2017: 1933–1942.

[113]HUANG M K, LI H, BAI B, et al. A federated multi-view deep learning framework for privacy-preserving recommendations[J]. arXiv preprint, 2020, arXiv: 2008.10808.

[114]HUANG P S, HE X D, GAO J F, et al.

Learning deep structured semantic models for web search using clickthrough data[C]//Proceedings of the 22nd ACM International Conference on Information & Knowledge Management. New York: ACM Press, 2013: 2333-2338.

[115]LIN Y J, REN P J, CHEN Z M, et al. Meta matrix factorization for federated rating predictions[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2020: 981-990.

[116]MUHAMMAD K, WANG Q Q, O'REILLY-MORGAN D, et al. FedFast: going beyond average for faster training of federated recommender systems[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2020: 1234-1242.

[117]KHARITONOV E. Federated online learning to rank with evolution strategies[C]//Proceedings of the 12th ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2019: 249-257.

[118]TRIENES J, CANO A T, HIEMSTRA D. Recommending users: whom to follow on federated social networks[J]. arXiv preprint, 2018, arXiv:181109292.

[119]TAN B, LIU B, ZHENG V, et al. A federated recommender system for online services[C]//Proceedings of the 14th ACM Conference on Recommender Systems. New York: ACM Press, 2020: 579-581.

[120]RIBERO M, HENDERSON J, WILLIAMSON S, et al. Federating recommendations using differentially private prototypes[J]. arXiv preprint, 2020, arXiv:200300602.

[121]HU H S, DOBBIE G, SALCIC Z, et al. A locality sensitive hashing based approach for federated recommender system[C]//Proceedings of 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing. Piscataway: IEEE Press, 2020: 836-842.

[122]BLANCHARD P, MHAMDI E M E, GUERRAOUI R, et al. Machine learning with adversaries: Byzantine tolerant gradient descent[C]//Proceedings of the 31st Annual Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 118-128.

[123]MHAMDI E M E, GUERRAOUI R, ROUAULT S E B. The hidden vulnerability of distributed learning in Byzantium[C]//Proceedings of the 34th International Conference on Machine Learning. New York: ACM Press, 2018.

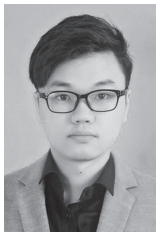
[124]WANG H, YUROCHKIN M, SUN Y, et al. Federated learning with matched averaging[J]. arXiv preprint, 2020, arXiv:2002.06440.

[125]YAN B J, LIU B Y, WANG L J, et al. FedCM: a real-time contribution measurement method for participants in federated learning[C]//Proceedings of 2021 International Joint Conference on Neural Networks. Piscataway: IEEE Press, 2021: 1-8.

[126]WANG G, DANG C X, ZHOU Z Y. Measure contribution of participants in federated learning[C]//Proceedings of 2019 IEEE International Conference on Big Data. Piscataway: IEEE Press, 2019: 2597-2604.

[127]ZHANG J F, LI C, ROBLES-KELLY A, et al. Hierarchically fair federated learning[J]. arXiv preprint, 2020, arXiv:2004.10386.

作者简介



朱智韬（1996– ），男，中国科学技术大学硕士生，平安科技（深圳）有限公司算法工程师，中国计算机学会（CCF）会员，主要研究方向为人工智能、联邦学习和推荐系统等。

司世景（1988– ），男，博士，平安科技（深圳）有限公司资深算法研究员，中国科学技术大学硕士生企业导师，CCF会员。发表机器学习、大数据和人工智能领域国际核心论文20余篇。

王健宗（1983– ），男，博士，平安科技（深圳）有限公司副总工程师、资深人工智能总监。CCF理事、杰出会员，CCF大数据专家委员会委员，主要研究方向为联邦学习、深度学习、云计算、物联网和元宇宙。

肖京（1972– ），男，博士，平安科技（深圳）有限公司首席科学家，深圳市政协委员，中国计算机学会深圳会员活动中心副主席，清华大学、上海交通大学、同济大学、香港中文大学、深圳大学、上海纽约大学客座教授，长期从事人工智能与大数据分析挖掘相关领域研究工作，发表计算机图形学、自动驾驶、3D显示、医疗诊断、联邦学习等领域国际核心论文230余篇，授权专利220余项。

收稿日期： 2021-11-05

通信作者： 王健宗，jzwang@188.com

基金项目： 广东省重点领域研发计划“新一代人工智能”重大专项（No.2021B0101400003）

Foundation Item: The Key Research and Development Program of Guangdong Province (No.2021B0101400003)