$$p^{\pi} := \sum_{j=1}^{p} \delta(\pi(j) = \pi) \tag{35}$$

算法 4 给出了 FM 模型在 MCMC 下解决回归 i 的详细优化过程.若想将其扩展到二分类任务的解决,则需要将正态分布的 $\hat{y}$ 映射到伯努利分布.这样,MCMC 算法就能预测一个实例属于正类或者负类的概率.

算法 4. 马尔可夫蒙特卡洛算法.

输入:训练集 D_{tr},测试集 D_{te},迭代次数 T,隐因子维度 k,初始化参数 σ;

输出:测试集 D_{te} 上预测结果 $\hat{y}$.

初始化模型参数 $w_0 \leftarrow 0, \boldsymbol{w} \leftarrow (0,\ldots,0), \boldsymbol{V} \sim N(0,\sigma)$

for iter in Range $\{1,\ldots,T\}$ **do**

  在训练集 D_{tr} 上预测计算 $\hat{y}$

  计算残差 $e(\boldsymbol{x}, y \mid \Theta) \leftarrow y - \hat{y}$

  按照公式(34)抽样得到 α

  for $(\mu_{\pi}^{\cdot}, \lambda_{\pi}^{\cdot}) \in \Theta_H$ **do**

    按照公式(34)抽样得到 $\lambda_{\pi}^{\cdot}$ 和 $\mu_{\pi}^{\cdot}$

  end for

  按照公式(31)抽样得到 w_0

  更新残差 $e(\boldsymbol{x},y|\Theta)$

  for $i \in \{1,\ldots,p\}$ **do**

    按照公式(31)抽样得到 w_i

    更新残差 $e(\boldsymbol{x},y|\Theta)$

  end for

  for $f \in \{1,\ldots,k\}$ **do**

    for $i \in \{1,\ldots,p\}$ **do**

      按照公式(31)抽样得到 $v_{i,f}$

      更新残差 $e(\boldsymbol{x},y|\Theta)$

    end for

  end for

  测试集上计算预测值 $\hat{y}_{test}^{*}$

  $\hat{y}_{test} \leftarrow \hat{y}_{test} + \hat{y}_{test}^{*}$

end for

求测试集预测平均值 $\hat{y}_{test} \leftarrow \hat{y}_{test} / T$

通过对 MCMC 方法分析可知 MCMC 的正则项超参 Θ_H 是抽样选取的,为此引进了新的超先验参数 Θ_0.但是,超先验参数的数目要小于正则化项超参的数量.另外,更重要的一点在于:相比随机梯度下降和交替最小二乘算法,MCMC 对于超先验参数 Θ_0 的选择不敏感.即使该参数选择不好,FM 模型也能得到较好的结果.

3.5 4种优化算法比较

本节综述了 FM 模型常用的 4 种优化学习方法,分别是随机梯度下降法、交替最小二乘法、基于自适应正则的随机梯度下降法和马尔可夫蒙特卡洛法.下面我们从时间复杂度、空间复杂度、适用任务场景、超参数类型这 4 个方面对上述方法总结说明,见表 2.其中,$N_z(\boldsymbol{X})$表示训练集中所有不为 0 特征的总个数.由于每次迭代时参数的更新计算通过遍历一次训练集即可完成,因此 4 种优化方法在时间复杂度上是一致的.然而在存储复杂度上,各个优化算法所需的存储空间是不同的.两类随机梯度下降算法下的 FM 优化只需要存储模型中每一个参数的迭代更新值,故其存储复杂度为常数级.与上述梯度更新相比,交替最小二乘法和马尔可夫蒙特卡洛方法除去 $1+p(k+1)$大小内存用于参数的存储外,还需要 $O(nk)$的内存去存储预计算结果.在适用任务场景中,使用不

同损失函数或分布的 4 种优化方法都能胜任常见的回归或分类问题.在超参类型上,随机梯度下降需用户自定义的超参最多,包括学习速率、分布初始化参数和正则化项系数.基于自适应正则化的随机梯度下降和马尔可夫蒙特卡洛法将正则化项系数作为参数糅合到模型中进行共同学习,避免了最优正则化项系数的搜索过程,模型学习过程更加健壮.而除了随机梯度下降,其他 3 种优化方法都不存在学习速率的指定,也使得模型的优化更加稳定.

Table 2 Comparison of different optimization methods

表 2 不同优化方法对比

算法名称	时间复杂度	空间复杂度	适用回归	适用分类	超参类型
随机梯度下降	$O(kN_z(\boldsymbol{X}))$	$O(1)$	是	是	学习速率/正则项系数/分布参数
基于自适应正则的随机梯度下降	$O(kN_z(\boldsymbol{X}))$	$O(1)$	是	是	学习速率/分布参数
交替最小二乘	$O(kN_z(\boldsymbol{X}))$	$O(nk)$	是	是	正则项系数/分布参数
马尔可夫蒙特卡洛法	$O(kN_z(\boldsymbol{X}))$	$O(nk)$	是	是	超先验参数/分布参数

4 存在问题及未来研究方向

4.1 存在问题

在第 2 节给出了因子分解机模型在准确性提升和性能加速两个方面取得的研究进展,然而仍存在以下几点不足.

(1) 模型准确性方面

现有对 FM 模型准确性提升的工作基本从低阶到高阶交互、神经网络与 FM 的结合、好的交互特征选择、概率模型推导、凸优化模型、在线学习、层次信息引入以及具体场景分析等 8 个方面展开.在高阶交互方面,基于更高阶特征交互的 FM 模型固然能够进一步挖掘特征之间的相互关联,从而提升模型准确率,然而 FM 的高阶交互会导致模型可解释性进一步降低,并可能进一步放大噪音特征对模型准确性的影响.在神经网络与 FM 的结合方向,现有结合方案可分为两类:① 将标准 FM 模型作为神经网络的输入,以便于后续利用神经网络完成特征的更高阶非线性交互;② 同时使用标准 FM 模型和神经网络分别完成特征的低阶和高阶交互建模,未来可考虑与其他更多类型神经网络的结合以适应不同的应用场景.在交互特征选择领域,已有的工作多是基于标准 FM 模型展开的,存在可扩展性较低以及模型参数规模并未因为特征选择而减小等问题.针对 FM 模型的在线学习,现有工作基于凸 FM 模型展开,分别利用流行的凸优化学习算法在线条件梯度和 FTRL 完成模型的更新.然而这种单机在线学习算法已无法适用于大规模参数和数据集的场景.最后,在层次信息引入上,研究者从树状层次特征引入和不同类别特征交互使用不同隐因子向量/权重两个角度展开来提高模型准确性,但是依然存在拘泥于特定场景、模型规模指数级增加及无法适用于高阶特征交互建模等问题.

(2) 模型效率方面

现有针对 FM 模型的分布式扩展工作已有多种方案,然而挑战依然存在.按照分布式框架的不同,可分为 Map-Redce/Spark、参数服务器框架和环形分布式框架.基于 Map-Reduce/Spark 平台的扩展既有高阶 FM 模型,也有标准 FM 算法,最大的短板在于全局模型需存放于一个服务器节点且模型传输时通信开销巨大,导致其在模型规模庞大的真实生产环境中无法应用.基于参数服务器框架的扩展目前仅限于标准 FM 模型,相比 Map-Reduce/Spark 架构,通信开销更少,模型存储也不再限制于一个服务器节点.然而由于模型的高维性,多次迭代下通信开销依然巨大.在此背景下,基于环形分布式的二阶 FM 模型被提出,优势在于通信开销和频次进一步减小,但稍显劣势之处在于单个节点上存储的模型规模有所增加.从上述描述可以看出:标准 FM 模型的分布式扩展已趋于成熟,然而基于 FM 的高阶交互、特征选择、凸模型以及在线学习等多种模型变种的分布式扩展依然存在瓶颈,有待进一步完善.

4.2 未来研究方向

针对上述因子分解机模型研究中依然存在的问题,对其未来研究方向进行探讨.

(1) 模型准确性研究

在高阶模型可解释性方面:一方面,可借助特征选择算法降低高阶建模时所需的特征数目,从而降低交互阶数,并剔除噪音、冗余和不相关特征;另一方面,对所有特征的相关性进行分析,使用多个高阶 FM 模对每组特征完成交互建模,最终根据任务场景,利用加权平均或投票表决的方式进行预测.

在神经网络与 FM 模型的结合方面,现有研究神经网络部分都是基于传统神经网络或向传统神经网络中加入注意力机制,后续研究可从神经网络部分入手,考虑 FM 模型与其他高级神经网络类型如长短期记忆网络 LSTM 的结合,可用于在流式数据中挖掘用户的长期和短期兴趣,提升推荐结果质量.

在交互特征选择方面,研究目标大致可分为两个方向:① 考虑扩展基于标准 FM 的交互特征选择至高阶 FM 模型,与现有高阶交互研究相结合,不仅可以达到降低模型参数规模的目标,而且提高了模型准确性和可解释性;② 考虑采用现有流行方法高效的选择好的交互特征,如可利用神经网络基于原始数据完成特征的表示学习,利用强化学习思想选择.

在 FM 模型的在线学习方面,有两个基本问题需要考虑:其一,假设新进数据特征维度不变,即不存在新特征的加入,这种情况下,如何利用新进数据增量更新已有模型;其二,假设新进数据中存在新用户、新物品和新的上下文特征,那么如何基于已有模型较小调整甚至不改变的情况下学习得到新的模型.现有工作都是以特征维度不变为前提展开的,后续的研究方向可从 3 个方向入手:① 利用其他已有的在线学习方法完成 FM 模型及其变种的增量更新;② 扩展单机环境下 FM 模型的在线学习方法至分布式环境下,以适应真实生产环境下的大规模参数和数据集;③ 考虑新的特征加入条件下,利用矩阵运算达到基于已有模型的较小调整甚至不改变情况下更大规模新的模型的学习.

在层次信息引入方面,现有工作主要存在特定场景特殊分析、模型规模指数级增长以及不适用于高阶特征交互建模等短板.因此,未来可考虑提出一种通用的层次特征建模方法,使得 FM 模型能够利用这种层次的特征有效提高模型的表达能力,如结合神经网络和注意力机制,使得特征间的高阶交互建模得以实现,并利用注意力机制学习得到不同类别特征交互的强度.

(2) 模型效率研究

现有针对 FM 模型效率提升主要分为基于数据/参数重组和基于分布式两种.其中,由于分布式优化在大数据环境下的高效性使其成为主流研究方向,得到较多关注.然而多数工作都是基于标准 FM 模型展开的,在变种 FM 模型,如高阶交互、特征选择等方面仅仅做了简单的分布式实现.在 FM 模型的增量更新方面,目前仍无相关研究.因此,如何针对上述两个方向进行分布式扩展是非常值得关注的.

5 总 结

作为一种通用的分解模型,因子分解机模型能够取得较好的预测和推荐结果,近年来在机器学习领域得到了广泛的应用和关注,取得了诸多研究成果.本文首先从分解模型的演化角度说明了传统的矩阵分解模型如何一步步进化到基于特定上下文的分解方法,进而得到本文综述重点——通用因子分解机模型,并通过因子分解机模型与其他流行的特定分解模型的相互关联性说明了因子分解机模型强大的泛化性;然后,从因子分解机模型的准确性和性能两方面出发,说明标准因子分解机模型存在的不足,并给出近年来研究者针对这两个方面所存在问题的优化方案;接着,从独立于问题模型的角度综述了 FM 模型常用的 4 种优化方法,并指出各个优化算法的优势和不足;最后指出了现有因子分解模型研究中存在的不足和未来可能的研究方向.

References:

[1] Gantz J, Reinsel D. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. IDC iView: IDC Analyze the Future, 2012,2007(2012):1−16.

[2] Huang LW, Jiang BT, Lv SY, Liu YB, Li DY. Survey on deep learning based recommender systems. Chinese Journal of Computers, 2018,41(7):1619−1647 (in Chinese with English abstract).

[3] Rendle S. Factorization machines. In: Proc. of the 2010 IEEE 10th Int’l Conf. on Data Mining (ICDM). IEEE, 2010. 995−1000.

[4] Koren Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In: Proc. of the ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. ACM Press, 2008. 426−434.

[5] Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized Markov chains for next-basket recommendation. In: Proc. of the Int'l Conf. on World Wide Web. ACM Press, 2010. 811−820.

[6] Koren Y. Collaborative filtering with temporal dynamics. Communications of the ACM, 2010,53(4):89−97.

[7] Xiong L, Chen X, Huang TK, *et al*. Temporal collaborative filtering with bayesian probabilistic tensor factorization. In: Proc. of the 2010 SIAM Int'l Conf. on Data Mining. 2010. 211−222.

[8] Rendle S, Schmidt-Thieme L. Pairwise interaction tensor factorization for personalized tag recommendation. In: Proc. of the ACM Int'l Conf. on Web Search and Data Mining. ACM Press, 2010. 81−90.

[9] Meng XW, Ji WY, Zhang YJ. A survey of recommendation systems in big data. Journal of Beijing University of Posts and Telecommunications, 2015,38(2):1−15 (in Chinese with English abstract).

[10] Dror G, Koenigstein N, Koren Y, *et al*. The Yahoo! music dataset and KDD-Cup'11. In: Proc. of the 17th ACM SIGKDD Conf. on Knowledge Discovery and Data Mining. San Diego: ACM Press, 2012. 8−18.

[11] Golub G, Kahan K. Calculating the singular values and pseudo-inverse of a matrix. Journal of the Society for Industrial and Applied Mathematics, 1965,2(2):205−224.

[12] Lee DD, Seung H. Algorithms for non-negative matrix factorization. In: Proc. of the 13th Advances in Neural Information Processing Systems (NIPS 2000). Denver: MIT Press, 2000. 556−562.

[13] Salakhutdinov R, Mnih A. Probabilistic matrix factorization. In: Proc. of the 20th Advances in Neural Information Processing Systems. 2007,20(3):432−451.

[14] Ding WF, Zheng XL, Chen DR. Active sampling based on PureSVD model for collaborative filtering. Journal of Beijing University of Posts and Telecommunications, 2013,36(4):23−26 (in Chinese with English abstract).

[15] Adomavicius G, Sankaranarayanan R, Sen S, *et al*. Incorporating contextual information in recommender systems using a multidimensional approach. ACM Trans. on Information Systems (TOIS), 2005,23(1):103−145.

[16] Symeonidis P, Papadimitriou A, Manolopoulos Y, *et al*. Geo-social recommendations based on incremental tensor reduction and local path traversal. In: Proc. of the 3rd ACM SIGSPATIAL Int'l Workshop on Location-Based Social Networks. 2011. 89−96.

[17] Tu DD, Shu CC, Yu HY. Using unified probabilistic matrix factorization for contextual advertisement recommendation. Ruan Jian Xue Bao/Journal of Software, 2013,24(3):454−464 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4238.htm [doi: 10.3724/SP.J.1001.2013.04238]

[18] Koren Y. Factor in the neighbors: Scalable and accurate collaborative filtering. ACM Trans. on Knowledge Discovery from Data, 2010,4(1):2010.

[19] Gantner Z, Drumond L, Freudenthaler C, *et al*. Learning attribute-to-feature mappings for cold-start recommendations. In: Proc. of the 2010 IEEE 10th Int'l Conf. on Data Mining (ICDM). IEEE, 2010. 176−185.

[20] Agarwal D, Chen BC. Regression-based latent factor models. In: Proc. of the 15th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. ACM Press, 2009. 19−28.

[21] Rendle S, Gantner Z, Freudenthaler C, *et al*. Fast context-aware recommendations with factorization machines. In: Proc. of the 34th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM Press, 2011. 635−644.

[22] Freudenthaler C, Schmidt-Thieme L, Rendle S. Bayesian factorization machines. 2011.

[23] Rendle S. Learning recommender systems with adaptive regularization. In: Proc. of the 5th ACM Int'l Conf. on Web Search and Data Mining. ACM Press, 2012. 133−142.

[24] Knoll J. Recommending with higher-order factorization machines. In: Proc. of the Research and Development in Intelligent Systems XXXIII. Springer Int'l Publishing, 2016. 103−116.

[25] Blondel M, Ishihata M, Fujino A, *et al*. Polynomial networks and factorization machines: New insights and efficient training algorithms. IEEE Trans. on Wireless Communications, 2016,15(1):131−145.

[26] Blondel M, Fujino A, Ueda N, *et al*. Higher-order factorization machines. Advances in Neural Information Processing Systems. 2016. 3351−3359.

[27] Lu CT, He L, Shao W, *et al*. Multilinear factorization machines for multi-task multi-view learning. In: Proc. of the 10th ACM Int'l Conf. on Web Search and Data Mining. ACM Press, 2017. 701−709.

[28] Lu CT, He L, Ding H, *et al*. Learning from multi-view multi-way data via structural factorization machines. In: Proc. of the WWW. 2018.

[29] Cao B, Zhou H, Li G, *et al*. Multi-view machines. In: Proc. of the Ninth ACM Int'l Conf. on Web Search and Data Mining. ACM, 2016. 427−436.

[30] Zheng L, Noroozi V, Yu PS. Joint deep modeling of users and items using reviews for recommendation. In: Proc. of the Int'l Conf. on Web Search and Data Mining. ACM Press, 2017. 425−434.

[31] Cai Y, Dong S, Hu J. Jointly modeling user and item reviews by CNN for multi-domain recommendation. In: Proc. of the China Conf. on Information Retrieval. Cham: Springer-Verlag, 2018. 237−248.

[32] Zhang W, Du T, Wang J. Deep learning over multi-field categorical data. In: Proc. of the European Conf. on Information Retrieval. Cham: Springer-Verlag, 2016. 45−57.

[33] Liu Y, Guo W, Zang D, *et al*. A hybrid neural network model with non-linear factorization machines for collaborative recommendation. In: Proc. of the China Conf. on Information Retrieval. Cham: Springer-Verlag, 2018. 213−224.

[34] Guo H, Tang R, Ye Y, *et al*. DeepFM: A factorization-machine based neural network for CTR prediction. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. AAAI Press, 2017. 1725−1731.

[35] He X, Chua TS. Neural factorization machines for sparse predictive analytics. In: Proc. of the 40th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM Press, 2017. 355−364.

[36] Xiao J, Ye H, He X, *et al*. Attentional factorization machines: Learning the weight of feature interactions via attention networks. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. AAAI Press, 2017. 3119−3125.

[37] Cheng C, Xia F, Zhang T, *et al*. Gradient boosting factorization machines. In: Proc. of the 8th ACM Conf. on Recommender Systems. ACM Press, 2014. 265−272.

[38] Xu J, Lin K, Tan PN, *et al*. Synergies that matter: Efficient interaction selection via sparse factorization machine. In: Proc. of the 2016 SIAM Int'l Conf. on Data Mining. Society for Industrial and Applied Mathematics, 2016. 108−116.

[39] Yurochkin M, Nguyen XL. Multi-way interacting regression via factorization machines. In: Proc. of the Advances in Neural Information Processing Systems. 2017. 2598−2606.

[40] Nguyen TV, Karatzoglou A, Baltrunas L. Gaussian process factorization machines for context-aware recommendations. In: Proc. of the 37th Int'l ACM SIGIR Conf. on Research & Development in Information Retrieval. ACM Press, 2014. 63−72.

[41] Saha A, Acharya A, Ravindran B, *et al*. Nonparametric poisson factorization machine. In: Proc. of the IEEE Int'l Conf. on Data Mining (ICDM 2015). IEEE, 2015. 967−972.

[42] Pan Z, Chen E, Liu Q, *et al*. Sparse factorization machines for click-through rate prediction. In: Proc. of the 2016 IEEE 16th Int'l Conf. on Data Mining (ICDM). IEEE, 2016. 400−409.

[43] Blondel M, Fujino A, Ueda N. Convex factorization machines. In: Proc. of the Joint European Conf. on Machine Learning and Knowledge Discovery in Databases. Cham: Springer-Verlag, 2015. 19−35.

[44] Chang Y, *et al*. Convex factorization machine for toxicogenomics prediction. In: Proc. of the Int'l Conf. on Knowledge Discovery and Data Mining. ACM Press, 2017. 1215−1224.

[45] Lin X, Zhang W, Zhang M, *et al*. Online compact convexified factorization machine. In: Proc. of the Int'l Conf. on World Wide Web. ACM Press, 2018. 1633−1642.

[46] Luo L, Zhang W, Zhang Z, *et al*. Sketched follow-the-regularized-leader for online factorization machine. In: Proc. of the 24th Int'l Conf. on Knowledge Discovery & Data Mining. ACM Press, 2018. 1900−1909.

[47] Juan Y, Zhuang Y, Chin WS, *et al*. Field-aware factorization machines for CTR prediction. In: Proc. of the 10th ACM Conf. on Recommender Systems. ACM Press, 2016. 43−50.

[48] Pan J, Xu J, Ruiz AL, *et al*. Field-weighted factorization machines for click-through rate prediction in display advertising. In: Proc. of the 2018 World Wide Web Conf. on World Wide Web. 2018. 1349−1357.

[49] Oentaryo R, Lim E, Low J, Lo D, Finegold M. Predicting response in mobile advertising with hierarchical importance-aware factorization machine. In: Proc. of the 7th ACM Int'l Conf. on Web Search and Data Mining. New York: ACM Press, 2014. 123−132.

[50] Hong L, Doumith A, Davison B. Co-factorization machines: Modeling user interests and predicting individual decisions in Twitter. In: Proc. of the 6th ACM Int'l Conf. on Web Search and Data Mining. ACM Press, 2013. 557−566.

[51] Loni B, Shi Y, Larson M, *et al*. Cross-domain collaborative filtering with factorization machines. In: Proc. of the European Conf. on Information Retrieval. Cham: Springer-Verlag, 2014. 656−661.

[52] Zhong E, Fan W, Yang Q. Contextual collaborative filtering via hierarchical matrix factorization. In: Proc. of the 2012 SIAM Int'l Conf. on Data Mining. Society for Industrial and Applied Mathematics, 2012. 744−755.

[53] Wang S, Du C, Zhao K, *et al*. Random partition factorization machines for context-aware recommendations. In: Proc. of the Int'l Conf. on Web-Age Information Management. Cham: Springer-Verlag, 2016. 219−230.

[54] Rendle S. Social network and click-through prediction with factorization machines. In: Proc. of the KDD-Cup Workshop. 2012. 113.

[55] Ding Y, Wang D, Xin X, *et al*. SCFM: Social and crowdsourcing factorization machines for recommendation. Applied Soft Computing, 2018,66:548−556.

[56] Zhou J, Wang D, Ding Y, *et al*. SocialFM: A social recommender system with factorization machines. In: Proc. of the Int'l Conf. on Web-Age Information Management. Cham: Springer-Verlag, 2016. 286−297.

[57] Qiang RW, Liang F, Yang JW. Exploiting ranking factorization machines for microblog retrieval. In: Proc. of the 22nd ACM Int'l Conf. on Information and Knowledge Management (CIKMn 2013). 2013. 1783−1788.

[58] Rendle S, Freudenthaler C, Gantner Z, *et al*. BPR: Bayesian personalized ranking from implicit feedback. In: Proc. of the 25th Conf. on Uncertainty in Artificial Intelligence. AUAI Press, 2009. 452−461.

[59] Guo W, Wu S, Wang L, *et al*. Personalized ranking with pairwise factorization machines. Neurocomputing, 2016,214:191−200.

[60] Yuan FJ, Guo GB, Jose JM, Chen L, Yu HT, Zhang WN. Lambdafm: Learning optimal ranking with factorization machines using lambda surrogates. In: Proc. of the 25th ACM Int'l on Conf. on Information and Knowledge Management (CIKM 2016). 2016. 227−236.

[61] Yuan FJ, Guo GB, Jose JM, *et al*. Boostfm: Boosted factorization machines for top-*n* feature-based recommendation. In: Proc. of the 22nd Int'l Conf. on Intelligent User Interfaces. ACM Press, 2017. 45−54.

[62] Rendle S. Scaling factorization machines to relational data. Proc. of the VLDB Endowment, 2013,6(5):337−348.

[63] Liu H, He X, Feng F, *et al*. Discrete factorization machines for fast feature-based recommendation. In: Proc. of the Int'l Joint Conf. on Artificial Intelligence. 2018. 3449−3455.

[64] Sun H, Wang W, Shi Z. Parallel factorization machine recommended algorithm based on MapReduce. In: Proc. of the Int'l Conf. on Semantics, Knowledge and Grids. IEEE, 2014. 120−123.

[65] Li M, Liu Z, Smola AJ, *et al*. Difacto: Distributed factorization machines. In: Proc. of the 9th ACM Int'l Conf. on Web Search and Data Mining. ACM Press, 2016. 377−386.

[66] Zhong E, Shi Y, Liu N, *et al*. Scaling factorization machines with parameter server. In: Proc. of the 25th ACM Int'l Conf. on Information and Knowledge Management. ACM Press, 2016. 1583−1592.

[67] Zhao K, Zhang J, Zhang L, *et al*. CDSFM: A circular distributed SGLD-based factorization machines. In: Proc. of the Int'l Conf. on Database Systems for Advanced Applications. Cham: Springer-Verlag, 2018. 701−709.

[68] Zhao KK, Zhang J, Zhang LF, *et al*. Signed network prediction method based on the client-to-client distributed framework. Ruan Jian Xue Bao/Journal of Software, 2018,29(3):614−626 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5447.htm [doi: 10.13328/j.cnki.jos.005447]

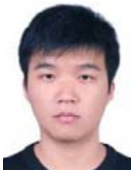
[69] Ma C, Liao Y, Wang Y, *et al*. F2M: Scalable field-aware factorization machines. In: Proc. of the MLSys on NIPS. 2016.

附中文参考文献:

[2] 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.计算机学报,2018,41(7):1619−1647.

[9] 孟祥武,纪威宇,张玉洁.大数据环境下的推荐系统.北京邮电大学学报,2015,38(2):1–15.

[14] 丁伟峰,郑小林,陈德人.基于 PureSVD 模型的协同过滤主动采样.北京邮电大学学报,2013,36(4):23−26.
[17] 涂丹丹,舒承椿,余海燕.基于联合概率矩阵分解的上下文广告推荐算法.软件学报,2013,24(3):454−464. http://www.jos.org.cn/1000-9825/4238.htm [doi: 10.3724/SP.J.1001.2013.04238]
[68] 赵衎衎,张静,张良富,等.基于端到端分布式框架的符号网络预测方法.软件学报,2018,29(3):614−626. http://www.jos.org.cn/1000-9825/5447.htm [doi: 10.13328/j.cnki.jos.005447]


赵衎衎(1991－),男,陕西渭南人,博士生,主要研究领域为推荐系统,大数据分析.


李翠平(1971－),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为社交网络分析,社会推荐,大数据分析及挖掘.


张良富(1991－),男,博士生,主要研究领域为机器学习,数据挖掘.


陈红(1965－),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为数据库技术,新硬件平台下的高性能计算.


张静(1984－),女,博士,讲师,CCF 专业会员,主要研究领域为数据挖掘,社会网络挖掘.