Workshop on Learning and Evaluating Recommendations with Impressions (LERI)

Maurizio Ferrari Dacrema
Politecnico di Milano
Milano, Italy
maurizio.ferrari@polimi.it

Pablo Castells
Universidad Autónoma de Madrid
Madrid, Spain
Amazon
Madrid, Spain
pablo.castells@uam.es

Justin Basilico
Netflix
Los Gatos, CA, USA
jbasilico@netflix.com

Paolo Cremonesi
Politecnico di Milano
Milano, Italy
ContentWise
Milano, Italy
paolo.cremonesi@polimi.it

ABSTRACT

Recommender systems typically rely on past user interactions as the primary source of information for making predictions. However, although highly informative, past user interactions are strongly biased. Impressions, on the other hand, are a new source of information that indicate the items displayed on screen when the user interacted (or not) with them, and have the potential to impact the field of recommender systems in several ways. Early research on impressions was constrained by the limited availability of public datasets, but this is rapidly changing and, as a consequence, interest in impressions has increased. Impressions present new research questions and opportunities, but also bring new challenges. Several works propose to use impressions as part of recommender models in various ways and discuss their information content. Others explore their potential in off-policy-estimation and reinforcement learning. Overall, the interest of the community is growing, but efforts in this direction remain disconnected. Therefore, we believe that a workshop would be useful in bringing the community together.

CCS CONCEPTS

• Information systems → Recommender systems; **Evaluation of retrieval results**; **• Computing methodologies → Reinforcement learning**; **• Human-centered computing → HCI design and evaluation methods**.

KEYWORDS

Recommender Systems, Impressions, Evaluation, User Modeling

ACM Reference Format:
Maurizio Ferrari Dacrema, Pablo Castells, Justin Basilico, and Paolo Cremonesi. 2023. Workshop on Learning and Evaluating Recommendations with Impressions (LERI). In *Seventeenth ACM Conference on Recommender Systems (RecSys '23), September 18–22, 2023, Singapore, Singapore.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3604915.3608756

RecSys '23, September 18–22, 2023, Singapore, Singapore

ACM ISBN 979-8-4007-0241-9/23/09.
https://doi.org/10.1145/3604915.3608756

1 MOTIVATION

In the early days of research on recommender systems, predictions were primarily based on past user interactions and user or item features. However, with advancements in technology, the scope and complexity of recommender systems have increased and new sources of data (such as context, knowledge-bases, and sequence structure) have emerged, driving the field forward and creating thriving sub-fields. Nevertheless, past user interactions remain the most potent and comprehensive source of predictive power. Despite this, observed interactions are a sparse and strongly biased source of information, which has significant implications for both learning from user actions and evaluating the quality of recommendations offline [26].

Recently, a source of information that was previously almost unavailable to the wider research community has emerged with the potential to impact the field in numerous ways: *impressions*. Impressions [7, 15, 25, 28, 37] refer to the items displayed on the screen when a user interacts (or not) with them and are the product of the whole recommendation engine [7, 21, 22]. Impressions constitute a nuanced and intricate data source that raises novel research questions, opportunities, and challenges. These may have profound implications for how recommender systems are conceptualized, trained, and evaluated.

Impressions took longer than ratings and interactions to cross the corporate boundary towards wider research availability [3, 11, 30, 34]. This started to happen eventually: early examples include the ACM RecSys Challenge in 2016, 2017 and 2019 [1, 2, 13], where the released datasets included impression data. Until recently, research was still limited by the lack of datasets, this was because the datasets released as part of the RecSys challenges are usually non redistributable and focus on very specific and narrow applications, while only very few other datasets were publicly available. This is rapidly changing and most of the available datasets including impressions have been published in the last few years: e.g., ContentWise Impressions [22], MIND [35], FINN.no Slates [7], Yahoo! [9] , Search Ads[1], Pandor [29], Ali-CCP [19], Alimama [27], Cross-shop Combo [40], In-Shop Combo [40], Kwai FAIR System [32], Kwai FAIR Experiment [32]. With the emergence of these new datasets, studying the use of impressions has become a more accessible topic for research. However, despite the increasing research interest, the efforts devoted to studying the use of impressions are still limited and fragmented. Therefore, we believe that organizing a workshop at this point would be both useful and timely in bringing together and consolidating the community working on this topic.

1.1 Status of Research, Challenges and Opportunities

Some works have already tried to use impressions to build better recommendation models in various ways: [4, 5, 16, 33, 38, 39] use impression data to compute features, re-ranking, sampling and to learn biases. Furthermore [6, 10, 14, 17, 18, 31, 36] apply neural or deep-learning models including impressions. Most of these papers have been published in the last two years in conferences such as SIGIR, KDD, WWW and RecSys.

Among the new research opportunities opened by impressions, being able to distinguish between the items that the user *observed* and *did not observe* could allow to provide better assumptions on how to label missing interactions. Some studies consider impressions to be a positive interaction signal, while others view them as negative signals [21].

Impressions also provide a direction for research that could help to bridge the gap between algorithms and user experience, two sides of recommender systems that are often studied independently of each other. For instance, continuously recommending the same item may lead to *user fatigue* [33], resulting in reduced user satisfaction with the system and wasted recommendations. By using impressions, recommender systems can better understand how users interact with the system and, thus, provide recommendations that improve user experience and engagement.

A further direction of research is in the evaluation of recommendation models. It is known that the past user interactions are a highly biased data source [26] and impressions, which represent the real behavior of the recommendation engine that acts as the intermediary between the user and the available catalogue, could allow to better identify those biases. The community is also exploring new methods for the evaluation of recommender systems, such as off-policy estimation (OPE) [8, 12, 23] and simulation environments [24] some of which already use impressions [20].

1.2 Workshop Description

The *Workshop on Learning and Evaluating Recommendations with Impressions* will focus on all aspects related to leveraging impression data to build and evaluate a recommendation engine. The goal is to both help to coalesce researchers exploring the use of impressions from different perspectives, as well as foster increased interest from the community for this new and still largely underexplored topic that has the potential of impacting the field in several ways. The workshop aims to provide a venue for researchers and practitioners to come together in order to: (i) share experience and lessons learned; (ii) identify key challenges in the area; (iii) build a common mental model and conceptual framework for thinking and researching on the use of impressions; (iv) identify emerging topics and new opportunities. The workshop also aims to lay bridges between practitioners and academics, encourage a wider availability of impression data sources and leverage industry's experience to guide and inform academic research.

1.3 Workshop Topics

Conceptual framework: definition of "impression", role of impressions in the recommendation task definition, user action attribution to impressions, prediction and causation, closed vs. open loops;

Recommendation models: new learning approaches taking advantage of impression data, impressions in label data, loss functions, model topologies;

Model training: data preprocessing, sampling, partitioning, hyperparameter tuning with impressions;

Evaluation: evaluation methodology and metrics, impact on offline evaluation bias;

User modeling: new models considering user behavior in face of impressed items;

Reinforcement learning and off-policy estimation: offline vs. online setting, impressions in RL and OPE;

Datasets: collection of new datasets with impressions from different domains, user interfaces, applications;

User Studies: how the user behavior is impacted by the composition of impressions, impact of user fatigue, etc.;

Theory: theoretical aspects in the use of impressions for recommender systems, both in the development of new and improved recommender systems and in their evaluation;

Perspectives: new perspectives on existing problems that could benefit or just change by adding impressions as a new variable, as well as old challenges that can be now tackled from new angles, and new challenges that derive from the use of impressions.

1.4 Workshop Organization

The workshop will be organized by:

Maurizio Ferrari Dacrema: Professor at Politecnico di Milano. His research interests include recommender systems evaluation and quantum computing. He has been local organization chair at the 12th Italian Information Retrieval Workshop.[2]

Pablo Castells: Professor at Universidad Autónoma de Madrid (UAM) and Amazon scholar. His research interests include recommender systems evaluation, algorithmic and experimental bias, and beyond-accuracy perspectives. He has organized six RecSys workshops in areas such as evaluation and experimentation, novelty and diversity, and industry applications; as well as workshops and tutorials at SIGIR, WSDM and The Web Conference. He has served in the RecSys organizing committee in different roles including

[1] https://www.kaggle.com/competitions/kddcup2012-track2

[2] https://recsyspolimi.github.io/iir2022/

PC co-chair in 2016, and served as PC co-chair and general co-chair of SIGIR in 2021 and 2022 respectively.

Justin Basilico: Netflix. He has been an Industry co-chair at RecSys 2022 and 2023, he has coorganized the 2020 and 2021 International Workshop on Industrial Recommendation Systems at KDD, and the 2022 REVEAL workshop at RecSys. He also coorganizes the annual Netflix Personalization, Recommendation, and Search (PRS) workshop.

Paolo Cremonesi: Professor at Politecnico di Milano and co-Founder of ContentWise. His research interests include recommender systems and quantum computing. He has served in the organization of scientific meetings in different roles, including program chair of ACM iTVX in 2013, and general co-chair of ACM RecSys in 2016. He serves in the RecSys steering committee since 2017.

2 RELATED PRIOR WORKSHOPS

We are not aware of any prior workshop that focused on the topic of impressions itself. However, the use of impressions has been connected to the following other workshops:

Causality, Counterfactuals, Sequential Decision-Making & Reinforcement Learning for Recommender Systems (RecSys 2022) the workshop did not discuss primarily impressions but the topic of off-policy estimation is connected to the availability of information on the real user preferences and on the bias introduced by the recommendation engine which could be estimated using impressions.

ACM RecSys Challenge Workshop (RecSys 2019, 2017 and 2016) the workshop did not discuss primarily impressions but the data available during the challenge included impressions and therefore some of the papers described how the teams used them.

RecSys workshops on recommender systems evaluation: Evaluation has been a recurring workshop topic at RecSys: workshops such as RUE 2012, RepSys 2013, REDD 2014, SimuRec 2021 have focused on offline evaluation methodology, metrics, reproducibility, bias, and datasets, among many other important elements and issues in recommender system evaluation. As far as the proposers are aware (as co-organizers of these past workshops), impressions were not addressed or discussed in that scope so far.

3 PROGRAM COMMITTEE

The following is a list of the confirmed program committee: **Antonio Ferrara** (Politecnico di Bari), **Claudio Pomo** (Politecnico di Bari), **Daniele Malitesta** (Politecnico di Bari), **David Massimo** (Free University of Bolzano), **Fernando B. Pérez Maurera** (Politecnico di Milano), **Marco de Gemmis** (Università degli Studi di Bari Aldo Moro), **Marco Polignano** (Università degli Studi di Bari Aldo Moro), **Maurizio Ferrari Dacrema** (Politecnico di Milano), **Nicolò Felicioni** (Politecnico di Milano), **Olivier Jeunen** (ShareChat), **Paolo Cremonesi** (Politecnico di Milano), **Pengjie Ren** (Shandong University), **Vito Walter Anelli** (Politecnico di Bari), **Xin Xin** (Shandong University).

REFERENCES

[1] Fabian Abel, András A. Benczúr, Daniel Kohlsdorf, Martha A. Larson, and Róbert Pálovics. 2016. RecSys Challenge 2016: Job Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, September 15-19, 2016*, Shilad Sen, Werner Geyer, Jill Freyne, and Pablo Castells (Eds.). ACM, 425–426. https://doi.org/10.1145/2959100.2959207

[2] Fabian Abel, Yashar Deldjoo, Mehdi Elahi, and Daniel Kohlsdorf. 2017. RecSys Challenge 2017: Offline and Online Evaluation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems, RecSys 2017, Como, Italy, August 27-31, 2017*, Paolo Cremonesi, Francesco Ricci, Shlomo Berkovsky, and Alexander Tuzhilin (Eds.). ACM, 372–373. https://doi.org/10.1145/3109859.3109954

[3] Deepak Agarwal, Bee-Chung Chen, and Pradheep Elango. 2009. Spatio-temporal models for estimating click-through rate. In *Proceedings of the 18th International Conference on World Wide Web, WWW 2009, Madrid, Spain, April 20-24, 2009*. ACM, 21–30. https://doi.org/10.1145/1526709.1526713

[4] Michal Aharon, Yohay Kaplan, Rina Levy, Oren Somekh, Ayelet Blanc, Neetai Eshel, Avi Shahar, Assaf Singer, and Alex Zlotnik. 2019. Soft Frequency Capping for Improved Ad Click Prediction in Yahoo Gemini Native. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, Wenwu Zhu, Dacheng Tao, Xueqi Cheng, Peng Cui, Elke A. Rundensteiner, David Carmel, Qi He, and Jeffrey Xu Yu (Eds.). ACM, 2793–2801. https://doi.org/10.1145/3357384.3357801

[5] Niv Buchbinder, Moran Feldman, Arpita Ghosh, and Joseph Naor. 2014. Frequency capping in online advertising. *J. Sched.* 17, 4 (2014), 385–398. https://doi.org/10.1007/s10951-014-0367-z

[6] Zhihong Chen, Jiawei Wu, Chenliang Li, Jingxu Chen, Rong Xiao, and Binqiang Zhao. 2022. Co-training Disentangled Domain Adaptation Network for Leveraging Popularity Bias in Recommenders. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, Enrique Amigó, Pablo Castells, Julio Gonzalo, Ben Carterette, J. Shane Culpepper, and Gabriella Kazai (Eds.). ACM, 60–69. https://doi.org/10.1145/3477495.3531952

[7] Simen Eide, David S. Leslie, and Arnoldo Frigessi. 2021. Dynamic Slate Recommendation with Gated Recurrent Units and Thompson Sampling. *CoRR* abs/2104.15046 (2021). arXiv:2104.15046 https://arxiv.org/abs/2104.15046

[8] Nicolò Felicioni, Maurizio Ferrari Dacrema, Marcello Restelli, and Paolo Cremonesi. 2022. Off-Policy Evaluation with Deficient Support Using Side Information. In *Advances in Neural Information Processing Systems (NeurIPS 2022)*. 15 pages.

[9] Claudio Gentile, Shuai Li, and Giovanni Zappella. 2014. Online Clustering of Bandits. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014 (JMLR Workshop and Conference Proceedings, Vol. 32)*. JMLR.org, 757–765. http://proceedings.mlr.press/v32/gentile14.html

[10] Shansan Gong and Kenny Q. Zhu. 2022. Positive, Negative and Neutral: Modeling Implicit Feedback in Session-based News Recommendation. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, Enrique Amigó, Pablo Castells, Julio Gonzalo, Ben Carterette, J. Shane Culpepper, and Gabriella Kazai (Eds.). ACM, 1185–1195. https://doi.org/10.1145/3477495.3532040

[11] Neha Gupta, Abhimanyu Das, Sandeep Pandey, and Vijay K. Narayanan. 2012. Factoring past exposure in display advertising targeting. In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '12, Beijing, China, August 12-16, 2012*, Qiang Yang, Deepak Agarwal, and Jian Pei (Eds.). ACM, 1204–1212. https://doi.org/10.1145/2339530.2339719

[12] Amir Hossein Jadidinejad, Craig Macdonald, and Iadh Ounis. 2022. The Simpson's Paradox in the Offline Evaluation of Recommendation Systems. *ACM Trans. Inf. Syst.* 40, 1 (2022), 4:1–4:22. https://doi.org/10.1145/3458509

[13] Peter Knees, Yashar Deldjoo, Farshad Bakhshandegan Moghaddam, Jens Adamczak, Gerard Paul Leyson, and Philipp Monreal. 2019. RecSys challenge 2019: session-based hotel recommendations. In *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019*, Toine Bogers, Alan Said, Peter Brusilovsky, and Domonkos Tikk (Eds.). ACM, 570–571. https://doi.org/10.1145/3298689.3346974

[14] Pengyang Li, Rong Chen, Quan Liu, Jian Xu, and Bo Zheng. 2022. Transform Cold-Start Users into Warm via Fused Behaviors in Large-Scale Recommendation. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, Enrique Amigó, Pablo Castells, Julio Gonzalo, Ben Carterette, J. Shane Culpepper, and Gabriella Kazai (Eds.). ACM, 2013–2017. https://doi.org/10.1145/3477495.3531797

[15] Dawen Liang, Laurent Charlin, James McInerney, and David M. Blei. 2016. Modeling User Exposure in Recommendation. In *Proceedings of the 25th International Conference on World Wide Web, WWW 2016, Montreal, Canada, April 11 - 15, 2016*, Jacqueline Bourdeau, Jim Hendler, Roger Nkambou, Ian Horrocks, and Ben Y. Zhao (Eds.). ACM, 951–961. https://doi.org/10.1145/2872427.2883090

[16] David C. Liu, Stephanie Kaye Rogers, Raymond Shiau, Dmitry Kislyuk, Kevin C. Ma, Zhigang Zhong, Jenny Liu, and Yushi Jing. 2017. Related Pins at Pinterest: The Evolution of a Real-World Recommender System. In *Proceedings of the 26th*

International Conference on World Wide Web Companion, Perth, Australia, April 3-7, 2017, Rick Barrett, Rick Cummings, Eugene Agichtein, and Evgeniy Gabrilovich (Eds.). ACM, 583–592. https://doi.org/10.1145/3041021.3054202

[17] Jiazhen Lou, Hong Wen, Fuyu Lv, Jing Zhang, Tengfei Yuan, and Zhao Li. 2022. Reweighting Negative Samples for Model-Agnostic Matching. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, Enrique Amigó, Pablo Castells, Julio Gonzalo, Ben Carterette, J. Shane Culpepper, and Gabriella Kazai (Eds.). ACM, 1823–1827. https://doi.org/10.1145/3477495.3532053

[18] Ning Ma, Mustafa Ispir, Yuan Li, Yongpeng Yang, Zhe Chen, Derek Zhiyuan Cheng, Lan Nie, and Kishor Barman. 2022. An Online Multi-task Learning Framework for Google Feed Ads Auction Models. In *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, Aidong Zhang and Huzefa Rangwala (Eds.). ACM, 3477–3485. https://doi.org/10.1145/3534678.3539055

[19] Xiao Ma, Liqin Zhao, Guan Huang, Zhi Wang, Zelin Hu, Xiaoqiang Zhu, and Kun Gai. 2018. Entire Space Multi-Task Model: An Effective Approach for Estimating Post-Click Conversion Rate. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR 2018, Ann Arbor, MI, USA, July 08-12, 2018*, Kevyn Collins-Thompson, Qiaozhu Mei, Brian D. Davison, Yiqun Liu, and Emine Yilmaz (Eds.). ACM, 1137–1140. https://doi.org/10.1145/3209978.3210104

[20] James McInerney, Ehtsham Elahi, Justin Basilico, Yves Raimond, and Tony Jebara. 2021. Accordion: A Trainable Simulator forLong-Term Interactive Systems. In *RecSys '21: Fifteenth ACM Conference on Recommender Systems, Amsterdam, The Netherlands, 27 September 2021 - 1 October 2021*, Humberto Jesús Corona Pampín, Martha A. Larson, Martijn C. Willemsen, Joseph A. Konstan, Julian J. McAuley, Jean Garcia-Gathright, Bouke Huurnink, and Even Oldridge (Eds.). ACM, 102–113. https://doi.org/10.1145/3460231.3474259

[21] Fernando B. Pérez Maurera, Maurizio Ferrari Dacrema, and Paolo Cremonesi. 2022. Towards the Evaluation of Recommender Systems with Impressions. In *RecSys '22: Sixteenth ACM Conference on Recommender Systems, Seattle, WA, USA, September 18 - 23, 2022*, Jennifer Golbeck, F. Maxwell Harper, Vanessa Murdock, Michael D. Ekstrand, Bracha Shapira, Justin Basilico, Keld T. Lundgaard, and Even Oldridge (Eds.). ACM, 610–615. https://doi.org/10.1145/3523227.3551483

[22] Fernando B. Pérez Maurera, Maurizio Ferrari Dacrema, Lorenzo Saule, Mario Scriminaci, and Paolo Cremonesi. 2020. ContentWise Impressions: An Industrial Dataset with Impressions Included. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, Mathieu d'Aquin, Stefan Dietze, Claudia Hauff, Edward Curry, and Philippe Cudré-Mauroux (Eds.). ACM, 3093–3100. https://doi.org/10.1145/3340531.3412774

[23] Doina Precup, Richard S. Sutton, and Satinder Singh. 2000. Eligibility Traces for Off-Policy Policy Evaluation. In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford University, Stanford, CA, USA, June 29 - July 2, 2000*, Pat Langley (Ed.). Morgan Kaufmann, 759–766.

[24] David Rohde, Stephen Bonner, Travis Dunlop, Flavian Vasile, and Alexandros Karatzoglou. 2018. RecoGym: A Reinforcement Learning Environment for the problem of Product Recommendation in Online Advertising. *CoRR* abs/1808.00720 (2018). arXiv:1808.00720 http://arxiv.org/abs/1808.00720

[25] Masahiro Sato, Janmajay Singh, Sho Takemori, Takashi Sonoda, Qian Zhang, and Tomoko Ohkuma. 2019. Uplift-based evaluation and optimization of recommenders. In *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019*, Toine Bogers, Alan Said, Peter Brusilovsky, and Domonkos Tikk (Eds.). ACM, 296–304. https://doi.org/10.1145/3298689.3347018

[26] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016 (JMLR Workshop and Conference Proceedings, Vol. 48)*, Maria-Florina Balcan and Kilian Q. Weinberger (Eds.). JMLR.org, 1670–1679. http://proceedings.mlr.press/v48/schnabel16.html

[27] Qijie Shen, Hong Wen, Wanjie Tao, Jing Zhang, Fuyu Lv, Zulong Chen, and Zhao Li. 2022. Deep Interest Highlight Network for Click-Through Rate Prediction in Trigger-Induced Recommendation. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, Frédérique Laforest, Raphaël Troncy, Elena Simperl, Deepak Agarwal, Aristides Gionis, Ivan Herman, and Lionel Médini (Eds.). ACM, 422–430. https://doi.org/10.1145/3485447.3511970

[28] Ting-Yi Shih, Ting-Chang Hou, Jian-De Jiang, Yen-Chieh Lien, Chia-Rui Lin, and Pu-Jen Cheng. 2016. Dynamically Integrating Item Exposure with Rating Prediction in Collaborative Filtering. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016*, Raffaele Perego, Fabrizio Sebastiani, Javed A. Aslam, Ian Ruthven, and Justin Zobel (Eds.). ACM, 813–816. https://doi.org/10.1145/2911451.2914769

[29] Sumit Sidana, Charlotte Laclau, and Massih-Reza Amini. 2018. Learning to recommend diverse items over implicit feedback on PANDOR. In *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys 2018, Vancouver, BC, Canada, October 2-7, 2018*, Sole Pera, Michael D. Ekstrand, Xavier Amatriain, and John O'Donovan (Eds.). ACM, 427–431. https://doi.org/10.1145/3240323.3240400

[30] Yu Song, Shuai Sun, Jianxun Lian and10.1145/3543846/non-impressions/reinforcement-learning-based-recommender-systems-a-survey Hong Huang, Yu Li, Hai Jin, and Xing Xie. 2022. Show Me the Whole World: Towards Entire Item Space Exploration for Interactive Personalized Recommendations. In *WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event / Tempe, AZ, USA, February 21 - 25, 2022*, K. Selcuk Candan, Huan Liu, Leman Akoglu, Xin Luna Dong, and Jiliang Tang (Eds.). ACM, 947–956. https://doi.org/10.1145/3488560.3498459

[31] Hao Wang, Tai-Wei Chang, Tianqiao Liu, Jianmin Huang, Zhichao Chen, Chao Yu, Ruopeng Li, and Wei Chu. 2022. ESCM2: Entire Space Counterfactual Multi-Task Model for Post-Click Conversion Rate Estimation. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, Enrique Amigó, Pablo Castells, Julio Gonzalo, Ben Carterette, J. Shane Culpepper, and Gabriella Kazai (Eds.). ACM, 363–372. https://doi.org/10.1145/3477495.3531972

[32] Jiayin Wang, Weizhi Ma, Jiayu Li, Hongyu Lu, Min Zhang, Biao Li, Yiqun Liu, Peng Jiang, and Shaoping Ma. 2022. Make Fairness More Fair: Fair Item Utility Estimation and Exposure Re-Distribution. In *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, Aidong Zhang and Huzefa Rangwala (Eds.). ACM, 1868–1877. https://doi.org/10.1145/3534678.3539354

[33] Chao-Yuan Wu, Christopher V. Alvino, Alexander J. Smola, and Justin Basilico. 2016. Using Navigation to Improve Recommendations in Real-Time. In *Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, September 15-19, 2016*, Shilad Sen, Werner Geyer, Jill Freyne, and Pablo Castells (Eds.). ACM, 341–348. https://doi.org/10.1145/2959100.2959174

[34] Chuhan Wu, Fangzhao Wu, Tao Qi, Chenliang Li, and Yongfeng Huang. 2022. Is News Recommendation a Sequential Recommendation Task?. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, Enrique Amigó, Pablo Castells, Julio Gonzalo, Ben Carterette, J. Shane Culpepper, and Gabriella Kazai (Eds.). ACM, 2382–2386. https://doi.org/10.1145/3477495.3531862

[35] Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, Chuhan Wu, Tao Qi, Jianxun Lian, Danyang Liu, Xing Xie, Jianfeng Gao, Winnie Wu, and Ming Zhou. 2020. MIND: A Large-scale Dataset for News Recommendation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetreault (Eds.). Association for Computational Linguistics, 3597–3606. https://doi.org/10.18653/v1/2020.acl-main.331

[36] Ruobing Xie, Shaoliang Zhang, Rui Wang, Feng Xia, and Leyu Lin. 2022. A Peep into the Future: Adversarial Future Encoding in Recommendation. In *WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event / Tempe, AZ, USA, February 21 - 25, 2022*, K. Selcuk Candan, Huan Liu, Leman Akoglu, Xin Luna Dong, and Jiliang Tang (Eds.). ACM, 1177–1185. https://doi.org/10.1145/3488560.3498476

[37] Xin Xin, Jiyuan Yang, Hanbing Wang, Jun Ma, Pengjie Ren, Hengliang Luo, Xinlei Shi, Zhumin Chen, and Zhaochun Ren. 2022. On the User Behavior Leakage from Recommender System Exposure. *ACM Trans. Inf. Syst.* (oct 2022). https://doi.org/10.1145/3568954 Just Accepted.

[38] XianXing Zhang, Yitong Zhou, Yiming Ma, Bee-Chung Chen, Liang Zhang, and Deepak Agarwal. 2016. GLMix: Generalized Linear Mixed Models For Large-Scale Response Prediction. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, Balaji Krishnapuram, Mohak Shah, Alexander J. Smola, Charu C. Aggarwal, Dou Shen, and Rajeev Rastogi (Eds.). ACM, 363–372. https://doi.org/10.1145/2939672.2939684

[39] Qian Zhao, Gediminas Adomavicius, F. Maxwell Harper, Martijn C. Willemsen, and Joseph A. Konstan. 2017. Toward Better Interactions in Recommender Systems: Cycling and Serpentining Approaches for Top-N Item Lists. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2017, Portland, OR, USA, February 25 - March 1, 2017*, Charlotte P. Lee, Steven E. Poltrock, Louise Barkhuus, Marcos Borges, and Wendy A. Kellogg (Eds.). ACM, 1444–1453. https://doi.org/10.1145/2998181.2998211

[40] Chenxu Zhu, Peng Du, Weinan Zhang, Yong Yu, and Yang Cao. 2022. Combo-Fashion: Fashion Clothes Matching CTR Prediction with Item History. In *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, Aidong Zhang and Huzefa Rangwala (Eds.). ACM, 4621–4629. https://doi.org/10.1145/3534678.3539101