

Workshop on Learning and Evaluating Recommendations with Impressions (LERI)*

Maurizio Ferrari Dacrema¹, Pablo Castells², Justin Basilico³ and Paolo Cremonesi^{1,4}

¹Politecnico di Milano, Italy

²Universidad Autónoma de Madrid, Spain and Amazon

³Netflix, Los Gatos, CA, USA

⁴ContentWise, Milano, Italy

Abstract

This volume contains the papers presented at the Workshop on Learning and Evaluating Recommendations with Impressions (LERI), held in conjunction with the 17th ACM Conference on Recommender Systems (RecSys 2023). 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, one of the aims of the LERI workshop is to bring the community together.

Keywords

Recommender Systems, Impressions, Evaluation, User Modeling

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 [1].

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 numer-

ous ways: *impressions*. Impressions [2, 3, 4, 5, 6] 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, 5, 8]. 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 [9, 10, 11, 12]. This started to happen eventually: early examples include the ACM RecSys Challenge in 2016, 2017 and 2019 [13, 14, 15], 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 [7], MIND [16], FINN.no Slates [5], Pandor [17], Ali-CCP [18], Alimama [19], Cross-shop Combo [20], In-Shop Combo [20], Kwai FAIR System [21], Kwai FAIR Experiment [21]. 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

Workshop on Learning and Evaluating Recommendations with Impressions (LERI) @ RecSys 2023, September 18-22 2023, Singapore

*Workshop held in conjunction with the 17th ACM Conference on Recommender Systems (RecSys 2023), in Singapore.

✉ maurizio.ferrari@polimi.it (M. Ferrari Dacrema);

pablo.castells@uam.es (P. Castells); jbasilico@netflix.com

(J. Basilico); paolo.cremonesi@polimi.it (P. Cremonesi)

📞 0000-0001-7103-2788 (M. Ferrari Dacrema); 0000-0003-0668-6317

(P. Castells); 0000-0002-3005-5200 (J. Basilico); 0000-0002-1253-8081

(P. Cremonesi)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License

Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

limited and fragmented. Therefore, one of the core aims of the LERI workshop is to bring together and consolidate 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: [22, 23, 24, 25, 26, 27] use impression data to compute features, re-ranking, sampling and to learn biases. Furthermore [28, 29, 30, 31, 32, 33, 34] 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 [8].

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* [25], 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 [1] 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) [35, 36, 37] and simulation environments [38] some of which already use impressions [39].

1.2. Workshop Description

The *Workshop on Learning and Evaluating Recommendations with Impressions* focuses 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 has been 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.¹

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 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.

The proceedings have been curated by:

Fernando Benjamín Pérez Maurera: Ph.D. candidate in Information Technology at Politecnico di Milano. His research interests include recommender systems evaluation and algorithmic design; specifically, impression-aware recommender systems.

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 program committee: **Antonio Ferrara** (Politecnico di Bari), **Claudio Pomo** (Politecnico di Bari), **Daniele Malitesta** (Politecnico di Bari), **David Massimo** (Free University of Bolzano), **Fernando Benjamín 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), **Pengjie Ren** (Shandong University), **Vito Walter Anelli** (Politecnico di Bari).

4. Timeline

The following is the timeline of LERI:

- Paper submission deadline: August 10th, 2023
- Author notification: August 28th, 2023
- Camera-ready version deadline: September 10th, 2023
- Workshop (at RecSys 2023): September 19th, 2023

¹<https://recsyspolimi.github.io/iir2022/>

5. Workshop Program

The workshop program was articulated as follows.²

Keynote by Jiangwei Pan: “Recommendation Modeling with Impression Data at Netflix”. The keynote discussed how impression data is used to build recommender models at Netflix, focusing in particular on which issues arise when using impressions to train a recommendation model at scale in a two-pass architecture, where impressions can be useful to perform a more fine-grained ranking. Furthermore, the keynote discussed how repeated impressions affect the user behavior, their corresponding long-term value and the issue of performing enough item exploration without negatively affecting the user satisfaction.³

Accepted Papers:

- [40] *Impression-Informed Multi-Behavior Recommender System: A Hierarchical Graph Attention Approach*; Dong Li; Divya Bhargavi; Vidya Sagar Ravipati
- [41] *Characterizing Impression-Aware Recommender Systems*; Fernando Benjamín Pérez Maurera; Maurizio Ferrari Dacrema; Pablo Castells; Paolo Cremonesi
- [42] *Effects of Human-curated Content on Diversity in PSM: ARD-M Dataset*; Marcel Hauck; Ahtsham Manzoor; Sven Pagel
- [43] *Formulating Video Watch Success Signals for Recommendations on Short Video Platforms*; Srijan Saket; Sai Baba Reddy Velugoti; Rishabh Mehrotra
- [44] *Offline Evaluation using Interactions to Decide Cross-selling Recommendations Algorithm for Online Food Delivery*; Manjit Madan
- [45] *Contextual Position Bias Estimation Using a Single Stochastic Logging Policy*; Giuseppe Di Benedetto; Alexander Buchholz; Ben London; Matej Jakimov; Yannik Stein; Jan Malte Lichtenberg; Vito Bellini; Matteo Ruffini; Thorsten Joachims
- [46] *Incorporating Impressions to Graph-Based Recommenders*; Fernando Benjamín Pérez Maurera; Maurizio Ferrari Dacrema; Pablo Castells; Paolo Cremonesi

Panel Discussion: Moderated by Paolo Cremonesi with panelists **Jiangwei Pan** (Netflix), **Arnab Bhadury** (YouTube) and **Srijan Saket** (Sharechat). The discussion focused on the challenges to conduct research on the use of impressions outside companies, due to the highly contextual nature of impression data and the strong connection of their information content to the specific domain and user interface. The panelists also discussed the usefulness of impressions in multimedia recommendation where it is difficult to obtain good metadata and content can have a short shelf-life, meaning impressions can be useful to provide more information in a shorter amount of time.

References

- [1] T. Schnabel, A. Swaminathan, A. Singh, N. Chandak, T. Joachims, Recommendations as treatments: Debiasing learning and evaluation, in: M. Balcan, K. Q. Weinberger (Eds.), Proceedings of the 33rd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016, volume 48 of *JMLR Workshop and Conference Proceedings*, JMLR.org, 2016, pp. 1670–1679. URL: <http://proceedings.mlr.press/v48/schnabel16.html>.
- [2] D. Liang, L. Charlin, J. McInerney, D. M. Blei, Modeling user exposure in recommendation, in: J. Bourdeau, J. Hendler, R. Nkambou, I. Horrocks, B. Y. Zhao (Eds.), Proceedings of the 25th International Conference on World Wide Web, WWW 2016, Montreal, Canada, April 11 - 15, 2016, ACM, 2016, pp. 951–961. URL: <https://doi.org/10.1145/2872427.2883090>. doi:10.1145/2872427.2883090.
- [3] T. Shih, T. Hou, J. Jiang, Y. Lien, C. Lin, P. Cheng, Dynamically integrating item exposure with rating prediction in collaborative filtering, in: R. Perego, F. Sebastiani, J. A. Aslam, I. Ruthven, J. Zobel (Eds.), Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016, ACM, 2016, pp. 813–816. URL: <https://doi.org/10.1145/2911451.2914769>. doi:10.1145/2911451.2914769.
- [4] M. Sato, J. Singh, S. Takemori, T. Sonoda, Q. Zhang, T. Ohkuma, Uplift-based evaluation and optimization of recommenders, in: T. Bogers, A. Said, P. Brusilovsky, D. Tikk (Eds.), Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019, ACM, 2019, pp. 296–304. URL: <https://doi.org/10.1145/3298689.3347018>. doi:10.1145/3298689.3347018.
- [5] S. Eide, D. S. Leslie, A. Frigessi, Dynamic slate recommendation with gated recurrent units

²<https://recsyspolimi.github.io/leri2023/>

³<https://www.slideshare.net/JiangweiPan/recommendation-modeling-with-impression-data-at-netflixpptx>

- and thompson sampling, CoRR abs/2104.15046 (2021). URL: <https://arxiv.org/abs/2104.15046>. arXiv:2104.15046.
- [6] X. Xin, J. Yang, H. Wang, J. Ma, P. Ren, H. Luo, X. Shi, Z. Chen, Z. Ren, On the user behavior leakage from recommender system exposure, *ACM Trans. Inf. Syst.* (2022). URL: <https://doi.org/10.1145/3568954>. doi:10.1145/3568954, just Accepted.
- [7] F. B. Pérez Maurera, M. Ferrari Dacrema, L. Saule, M. Scriminaci, P. Cremonesi, Contentwise impressions: An industrial dataset with impressions included, in: M. d’Aquin, S. Dietze, C. Hauff, E. Curry, P. Cudré-Mauroux (Eds.), *CIKM ’20: The 29th ACM International Conference on Information and Knowledge Management*, Virtual Event, Ireland, October 19-23, 2020, ACM, 2020, pp. 3093–3100. URL: <https://doi.org/10.1145/3340531.3412774>. doi:10.1145/3340531.3412774.
- [8] F. B. Pérez Maurera, M. Ferrari Dacrema, P. Cremonesi, Towards the evaluation of recommender systems with impressions, in: J. Golbeck, F. M. Harper, V. Murdock, M. D. Ekstrand, B. Shapira, J. Basilico, K. T. Lundgaard, E. Oldridge (Eds.), *RecSys ’22: Sixteenth ACM Conference on Recommender Systems*, Seattle, WA, USA, September 18 - 23, 2022, ACM, 2022, pp. 610–615. URL: <https://doi.org/10.1145/3523227.3551483>. doi:10.1145/3523227.3551483.
- [9] D. Agarwal, B. Chen, P. Elango, 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, 2009, pp. 21–30. URL: <https://doi.org/10.1145/1526709.1526713>. doi:10.1145/1526709.1526713.
- [10] N. Gupta, A. Das, S. Pandey, V. K. Narayanan, Factoring past exposure in display advertising targeting, in: Q. Yang, D. Agarwal, J. Pei (Eds.), *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’12*, Beijing, China, August 12-16, 2012, ACM, 2012, pp. 1204–1212. URL: <https://doi.org/10.1145/2339530.2339719>. doi:10.1145/2339530.2339719.
- [11] C. Wu, F. Wu, T. Qi, C. Li, Y. Huang, Is news recommendation a sequential recommendation task?, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), *SIGIR ’22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 2382–2386. URL: <https://doi.org/10.1145/3477495.3531862>. doi:10.1145/3477495.3531862.
- [12] Y. Song, S. Sun, J. L. and10.1145/3543846/non-impressions/reinforcement-learning-based-recommender-systems-a-survey Hong Huang, Y. Li, H. Jin, X. Xie, Show me the whole world: Towards entire item space exploration for interactive personalized recommendations, in: K. S. Candan, H. Liu, L. Akoglu, X. L. Dong, J. Tang (Eds.), *WSDM ’22: The Fifteenth ACM International Conference on Web Search and Data Mining*, Virtual Event / Tempe, AZ, USA, February 21 - 25, 2022, ACM, 2022, pp. 947–956. URL: <https://doi.org/10.1145/3488560.3498459>. doi:10.1145/3488560.3498459.
- [13] F. Abel, A. A. Benczúr, D. Kohlsdorf, M. A. Larson, R. Pálovics, Recsys challenge 2016: Job recommendations, in: S. Sen, W. Geyer, J. Freyne, P. Castells (Eds.), *Proceedings of the 10th ACM Conference on Recommender Systems*, Boston, MA, USA, September 15-19, 2016, ACM, 2016, pp. 425–426. URL: <https://doi.org/10.1145/2959100.2959207>. doi:10.1145/2959100.2959207.
- [14] F. Abel, Y. Deldjoo, M. Elahi, D. Kohlsdorf, Recsys challenge 2017: Offline and online evaluation, in: P. Cremonesi, F. Ricci, S. Berkovsky, A. Tuzhilin (Eds.), *Proceedings of the Eleventh ACM Conference on Recommender Systems, RecSys 2017*, Como, Italy, August 27-31, 2017, ACM, 2017, pp. 372–373. URL: <https://doi.org/10.1145/3109859.3109954>. doi:10.1145/3109859.3109954.
- [15] P. Knees, Y. Deldjoo, F. B. Moghaddam, J. Adamczak, G. P. Leyson, P. Monreal, Recsys challenge 2019: session-based hotel recommendations, in: T. Bogers, A. Said, P. Brusilovsky, D. Tikk (Eds.), *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019*, Copenhagen, Denmark, September 16-20, 2019, ACM, 2019, pp. 570–571. URL: <https://doi.org/10.1145/3298689.3346974>. doi:10.1145/3298689.3346974.
- [16] F. Wu, Y. Qiao, J. Chen, C. Wu, T. Qi, J. Lian, D. Liu, X. Xie, J. Gao, W. Wu, M. Zhou, MIND: A large-scale dataset for news recommendation, in: D. Jurafsky, J. Chai, N. Schluter, J. R. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020*, Online, July 5-10, 2020, Association for Computational Linguistics, 2020, pp. 3597–3606. URL: <https://doi.org/10.18653/v1/2020.acl-main.331>. doi:10.18653/v1/2020.acl-main.331.
- [17] S. Sidana, C. Laclau, M. Amini, Learning to recommend diverse items over implicit feedback on PANDOR, in: S. Pera, M. D. Ekstrand, X. Amatriain, J. O’Donovan (Eds.), *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys 2018*, Vancouver, BC, Canada, October 2-7, 2018, ACM, 2018, pp. 427–431. URL: <https://doi.org/10.1145/3240323.3240400>. doi:10.1145/3240323.3240400.
- [18] X. Ma, L. Zhao, G. Huang, Z. Wang, Z. Hu, X. Zhu,

- K. Gai, Entire space multi-task model: An effective approach for estimating post-click conversion rate, in: K. Collins-Thompson, Q. Mei, B. D. Davison, Y. Liu, E. Yilmaz (Eds.), The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR 2018, Ann Arbor, MI, USA, July 08-12, 2018, ACM, 2018, pp. 1137–1140. URL: <https://doi.org/10.1145/3209978.3210104>. doi:10.1145/3209978.3210104.
- [19] Q. Shen, H. Wen, W. Tao, J. Zhang, F. Lv, Z. Chen, Z. Li, Deep interest highlight network for click-through rate prediction in trigger-induced recommendation, in: F. Laforest, R. Troncy, E. Simperl, D. Agarwal, A. Gionis, I. Herman, L. Médini (Eds.), WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022, ACM, 2022, pp. 422–430. URL: <https://doi.org/10.1145/3485447.3511970>. doi:10.1145/3485447.3511970.
- [20] C. Zhu, P. Du, W. Zhang, Y. Yu, Y. Cao, Combofashion: Fashion clothes matching CTR prediction with item history, in: A. Zhang, H. Rangwala (Eds.), KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022, ACM, 2022, pp. 4621–4629. URL: <https://doi.org/10.1145/3534678.3539101>. doi:10.1145/3534678.3539101.
- [21] J. Wang, W. Ma, J. Li, H. Lu, M. Zhang, B. Li, Y. Liu, P. Jiang, S. Ma, Make fairness more fair: Fair item utility estimation and exposure redistribution, in: A. Zhang, H. Rangwala (Eds.), KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022, ACM, 2022, pp. 1868–1877. URL: <https://doi.org/10.1145/3534678.3539354>. doi:10.1145/3534678.3539354.
- [22] N. Buchbinder, M. Feldman, A. Ghosh, J. Naor, Frequency capping in online advertising, *J. Sched.* 17 (2014) 385–398. URL: <https://doi.org/10.1007/s10951-014-0367-z>. doi:10.1007/s10951-014-0367-z.
- [23] Q. Zhao, G. Adomavicius, F. M. Harper, M. C. Willemsen, J. A. Konstan, Toward better interactions in recommender systems: Cycling and serpentine approaches for top-n item lists, in: C. P. Lee, S. E. Poltrock, L. Barkhuus, M. Borges, W. A. Kellogg (Eds.), Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2017, Portland, OR, USA, February 25 - March 1, 2017, ACM, 2017, pp. 1444–1453. URL: <https://doi.org/10.1145/2998181.2998211>. doi:10.1145/2998181.2998211.
- [24] D. C. Liu, S. K. Rogers, R. Shiau, D. Kislyuk, K. C. Ma, Z. Zhong, J. Liu, Y. Jing, Related pins at pinterest: The evolution of a real-world recommender system, in: R. Barrett, R. Cummings, E. Agichtein, E. Gabrilovich (Eds.), Proceedings of the 26th International Conference on World Wide Web Companion, Perth, Australia, April 3-7, 2017, ACM, 2017, pp. 583–592. URL: <https://doi.org/10.1145/3041021.3054202>. doi:10.1145/3041021.3054202.
- [25] C. Wu, C. V. Alvino, A. J. Smola, J. Basilico, Using navigation to improve recommendations in real-time, in: S. Sen, W. Geyer, J. Freyne, P. Castells (Eds.), Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, September 15-19, 2016, ACM, 2016, pp. 341–348. URL: <https://doi.org/10.1145/2959100.2959174>. doi:10.1145/2959100.2959174.
- [26] X. Zhang, Y. Zhou, Y. Ma, B. Chen, L. Zhang, D. Agarwal, Glmix: Generalized linear mixed models for large-scale response prediction, in: B. Krishnapuram, M. Shah, A. J. Smola, C. C. Aggarwal, D. Shen, R. Rastogi (Eds.), Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016, ACM, 2016, pp. 363–372. URL: <https://doi.org/10.1145/2939672.2939684>. doi:10.1145/2939672.2939684.
- [27] M. Aharon, Y. Kaplan, R. Levy, O. Somekh, A. Blanc, N. Eshel, A. Shahar, A. Singer, A. Zlotnik, Soft frequency capping for improved ad click prediction in yahoo gemini native, in: W. Zhu, D. Tao, X. Cheng, P. Cui, E. A. Rundensteiner, D. Carmel, Q. He, J. X. Yu (Eds.), Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019, ACM, 2019, pp. 2793–2801. URL: <https://doi.org/10.1145/3357384.3357801>. doi:10.1145/3357384.3357801.
- [28] Z. Chen, J. Wu, C. Li, J. Chen, R. Xiao, B. Zhao, Co-training disentangled domain adaptation network for leveraging popularity bias in recommenders, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 60–69. URL: <https://doi.org/10.1145/3477495.3531952>. doi:10.1145/3477495.3531952.
- [29] P. Li, R. Chen, Q. Liu, J. Xu, B. Zheng, Transform cold-start users into warm via fused behaviors in large-scale recommendation, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 2013–2017. URL: <https://doi.org/10.1145/3477495.3531797>. doi:10.1145/3477495.3531797.
- [30] J. Lou, H. Wen, F. Lv, J. Zhang, T. Yuan, Z. Li,

- Re-weighting negative samples for model-agnostic matching, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 1823–1827. URL: <https://doi.org/10.1145/3477495.3532053>. doi:10.1145/3477495.3532053.
- [31] S. Gong, K. Q. Zhu, Positive, negative and neutral: Modeling implicit feedback in session-based news recommendation, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 1185–1195. URL: <https://doi.org/10.1145/3477495.3532040>. doi:10.1145/3477495.3532040.
- [32] H. Wang, T. Chang, T. Liu, J. Huang, Z. Chen, C. Yu, R. Li, W. Chu, ESCM2: entire space counterfactual multi-task model for post-click conversion rate estimation, in: E. Amigó, P. Castells, J. Gonzalo, B. Carterette, J. S. Culpepper, G. Kazai (Eds.), SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, ACM, 2022, pp. 363–372. URL: <https://doi.org/10.1145/3477495.3531972>. doi:10.1145/3477495.3531972.
- [33] N. Ma, M. Ispir, Y. Li, Y. Yang, Z. Chen, D. Z. Cheng, L. Nie, K. Barman, An online multi-task learning framework for google feed ads auction models, in: A. Zhang, H. Rangwala (Eds.), KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022, ACM, 2022, pp. 3477–3485. URL: <https://doi.org/10.1145/3534678.3539055>. doi:10.1145/3534678.3539055.
- [34] R. Xie, S. Zhang, R. Wang, F. Xia, L. Lin, A peep into the future: Adversarial future encoding in recommendation, in: K. S. Candan, H. Liu, L. Akoglu, X. L. Dong, J. Tang (Eds.), WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event / Tempe, AZ, USA, February 21 - 25, 2022, ACM, 2022, pp. 1177–1185. URL: <https://doi.org/10.1145/3488560.3498476>. doi:10.1145/3488560.3498476.
- [35] D. Precup, R. S. Sutton, S. Singh, Eligibility traces for off-policy policy evaluation, in: P. Langley (Ed.), Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford University, Stanford, CA, USA, June 29 - July 2, 2000, Morgan Kaufmann, 2000, pp. 759–766.
- [36] N. Felicioni, M. Ferrari Dacrema, M. Restelli, P. Cremonesi, Off-policy evaluation with deficient support using side information, in: Advances in Neural Information Processing Systems (NeurIPS 2022), 2022, p. 15 pages.
- [37] A. H. Jadidinejad, C. Macdonald, I. Ounis, The simpson's paradox in the offline evaluation of recommendation systems, ACM Trans. Inf. Syst. 40 (2022) 4:1–4:22. URL: <https://doi.org/10.1145/3458509>. doi:10.1145/3458509.
- [38] D. Rohde, S. Bonner, T. Dunlop, F. Vasile, A. Karatzoglou, Recogym: A reinforcement learning environment for the problem of product recommendation in online advertising, CoRR abs/1808.00720 (2018). URL: <http://arxiv.org/abs/1808.00720>. arXiv:1808.00720.
- [39] J. McInerney, E. Elahi, J. Basilico, Y. Raimond, T. Jebara, Accordion: A trainable simulator for long-term interactive systems, in: H. J. C. Pampin, M. A. Larson, M. C. Willemsen, J. A. Konstan, J. J. McAuley, J. Garcia-Gathright, B. Huurmink, E. Oldridge (Eds.), RecSys '21: Fifteenth ACM Conference on Recommender Systems, Amsterdam, The Netherlands, 27 September 2021 - 1 October 2021, ACM, 2021, pp. 102–113. URL: <https://doi.org/10.1145/3460231.3474259>. doi:10.1145/3460231.3474259.
- [40] D. Li, D. Bhargavi, V. S. Ravipati, Impression-informed multi-behavior recommender system: A hierarchical graph attention approach, in: Proceedings of the Workshop on Learning and Evaluating Recommendations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 9–21.
- [41] F. B. Pérez Maurera, M. Ferrari Dacrema, P. Castells, P. Cremonesi, Characterizing impression-aware recommender systems, in: Proceedings of the Workshop on Learning and Evaluating Recommendations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 22–33.
- [42] M. Hauck, A. Manzoor, S. Pagel, Effects of human-curated content on diversity in PSM: ARD-M dataset, in: Proceedings of the Workshop on Learning and Evaluating Recommendations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 34–40.
- [43] S. Saket, S. B. R. Velugoti, R. Mehrotra, Formulating video watch success signals for recommendations on short video platforms, in: Proceedings of the Workshop on Learning and Evaluating Recommen-

- dations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 41–48.
- [44] M. Madan, Offline evaluation using interactions to decide cross-selling recommendations algorithm for online food delivery, in: Proceedings of the Workshop on Learning and Evaluating Recommendations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 49–54.
- [45] G. Di Benedetto, A. Buchholz, B. London, M. Jakimov, Y. Stein, J. M. Lichtenberg, V. Bellini, M. Ruffini, T. Joachims, Contextual position bias estimation using a single stochastic logging policy, in: Proceedings of the Workshop on Learning and Evaluating Recommendations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 55–61.
- [46] F. B. Pérez Maurera, M. Ferrari Dacrema, P. Castells, P. Cremonesi, Incorporating impressions to graph-based recommenders, in: Proceedings of the Workshop on Learning and Evaluating Recommendations with Impressions (LERI) co-located with the 17th ACM Conference on Recommender Systems (RecSys 2023), Singapore, 18th-22nd September 2023, CEUR Workshop Proceedings, CEUR-WS.org, 2023, pp. 62–67.