KuaiRand: An Unbiased Sequential Recommendation Dataset with Randomly Exposed Videos

Chongming Gao*
chongming.gao@gmail.com
University of Science and Technology
of China

Shijun Li*
lishijun@mail.ustc.edu.cn
University of Science and Technology
of China

Yuan Zhang*
yuanz.pku@gmail.com
Kuaishou Technology Co., Ltd.
China

Jiawei Chen
sleepyhunt@zju.edu.cn
Zhejiang University

Biao Li
libiao@kuaishou.com
Kuaishou Technology Co., Ltd.

Wenqiang Lei
wenqianglei@gmail.com
Sichuan University

Peng Jiang
jiangpeng@kuaishou.com
Kuaishou Technology Co., Ltd.
China

Xiangnan He
xiangnanhe@gmail.com
University of Science and Technology
of China

ABSTRACT

Recommender systems deployed in real-world applications can have inherent exposure bias, which leads to the biased logged data plaguing the researchers. A fundamental way to address this thorny problem is to collect users' interactions on randomly expose items, i.e., the missing-at-random data. A few works have asked certain users to rate or select randomly recommended items, e.g., Yahoo! [16], Coat [22], and OpenBandit [19]. However, these datasets are either too small in size or lack key information, such as unique user ID or the features of users/items. In this work, we present *KuaiRand*, an unbiased sequential recommendation dataset containing millions of intervened interactions on randomly exposed videos, collected from the video-sharing mobile App, Kuaishou. Different from existing datasets, KuaiRand records 12 kinds of user feedback signals (e.g., click, like, and view time) on randomly exposed videos inserted in the recommendation feeds in two weeks. To facilitate model learning, we further collect rich features of users and items as well as users' behavior history. By releasing this dataset, we enable the research of advanced debiasing large-scale recommendation scenarios for the first time. Also, with its distinctive features, KuaiRand can support various other research directions such as interactive recommendation, long sequential behavior modeling, and multi-task learning. The dataset and its news will be available at https://kuairand.com.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.
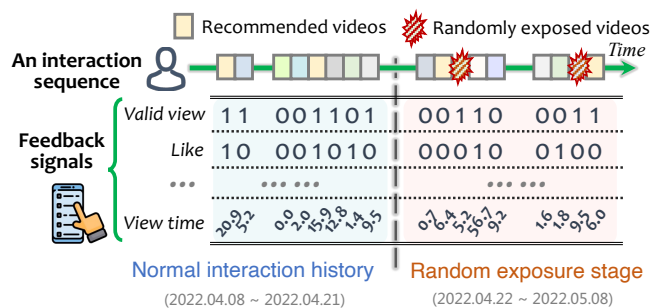
---

*Three authors contributed equally to this research.

Figure 1: Illustration of the proposed *KuaiRand* dataset.

KEYWORDS

Datasets; Recommendation; Random exposure; Long sequence

1 INTRODUCTION

Recommender systems have become a powerful tool for tech companies to connect users with appropriate items or services. Although developing and evaluating a recommendation model with real users are effective, online A/B testing is time- and money-consuming and entails the risk of hurting customer engagement, thus impairing the platform's value proposition [7, 21]. Hence, most of the efforts have been put into developing advanced recommendation models on the offline data generated from the interaction logs.

However, the bottleneck becomes the offline data *per se*. Since the highly sparse data comes from the collected historical interactions between users and the system, it is unavoidably affected by the exposure bias [3, 13, 14]. The biased data has catastrophic effects on researchers and practitioners of recommender systems. It leads to inconsistent online and offline results of the model [22] and

Table 1: Comparison of currently available datasets with randomly exposed data.

Dataset	Collection policy	# Users	# Items	# Interactions	# User features	# Item features	Timestamp	Feedback
Yahoo!R3 [16]	Normal interaction	15,400	1,000	311,704	0	0	No	Rating: [1,5]
	Random policy	5,400	1,000	54,000	7	0	No	Rating: [1,5]
Yahoo!R6A [11]	Random policy	No ID	271	45,811,883	6	6	Yes	Click
Yahoo!R6B [11]	Random policy	No ID	652	27,777,695	136	0	Yes	Click
Coat [22]	User self-selected	290	300	6,960	4	4	No	Rating: [1,5]
	Random policy	290	300	4,640	4	4	No	Rating: [1,5]
Open Bandit [19]	Bernoulli TS	No ID	80	24,011,308	4	4	Yes	Click
	Random policy	No ID	80	2,691,861	4	4	Yes	Click
KuaiRand-27K	15 policies	27,285	32,038,725	322,278,385	30	62	Yes	
	Random policy	27,285	7,583	1,186,059	30	62	Yes	12 signals
KuaiRand-1K	15 policies	1,000	4,369,953	11,713,045	30	62	Yes	(e.g., click,
	Random policy	1,000	7,388	43,028	30	62	Yes	like, hate,
KuaiRand-Pure	15 policies	27,285	7,551	1,436,609	30	62	Yes	view time)
	Random policy	27,285	7,583	1,186,059	30	62	Yes	

gives rise to filter bubbles [4], which can incur heavy losses for the companies and hinder the development of recommender systems.

1.1 Open Problem: Collecting Random Data

The fundamental problem lies in the fact that we have no knowledge about the massive missing interactions in the offline data [6]. To tackle the problem at the root, we have to collect the user preferences regarding those previously NOT exposed items. Gao et al. [6] collected a fully-observed dataset with each user's preferences on each item known. However, this dataset can inevitably be biased towards popular items due to its extremely high density.

A more practical way to obtain truly unbiased data is to collect users' feedback on uniformly sampled items, i.e., the missing-at-random (MAR) data [29]. It has been demonstrated that we can conduct an unbiased offline evaluation on the MAR data [20, 29].

Few works have considered this solution and collected user feedback on randomly exposed items. The well-known datasets Yahoo! [16] and Coat [22] asked a set of people to explicitly rate a certain number of randomly selected items in the score range from 1 to 5. However, the scales of the two datasets are quite limited — only 54, 000 ratings in Yahoo! and 4, 640 ratings in Coat for the randomly selected items. Besides, both of them lack some key information such as user/item features and timestamp. The recently introduced dataset OpenBandit [19] collected users' clicks on items from a slate of three items. Each slate is generated by a multi-armed bandit policy or a random policy. However, it only evolves 80 items, and the user identification (ID) information is removed. These insufficient information problems make existing datasets fail to satisfy the requirement of many recommendation tasks.

1.2 Contributions: The KuaiRand Dataset

Our goal is to enable unbiased offline evaluation in real-world recommendation scenarios. To this end, we present *KuaiRand*, a large-scale sequential recommendation dataset, collected from Kuaishou

App[1], one of the China's largest video-sharing Apps with more than 300 million daily active users. Different from previous datasets, it intervenes in the recommendation policies by randomly inserting randomly chosen videos in the normal recommendation feed for two weeks without user awareness. Hence, we can collect genuine unbiased user reflections on these randomly exposed videos. We recorded all 12 kinds of feedback signals, such as *click*, *adding to favorites*, and *view time*. To facilitate model learning, we further collected users' historical behaviors as well as rich side information of users/items. The process is illustrated in Fig. 1 and the detailed information can be referred to Table 1.

It is worth mentioning that KuaiRand is the first unbiased sequential recommendation data where we intervene in the original recommendation policy millions of times by inserting random items. The unbiased sequential data enables us to carry out the unbiased offline evaluation [20, 29], thus facilitating the research of debiasing [3] in large-scale recommendation scenarios. Furthermore, with its distinctive features, KuaiRand can naturally facilitate a variety of other recommendation tasks, such as interactive recommendation [4, 27], long sequential behavior modeling [17, 18], and multi-task learning [2, 28].

2 RELATED WORK

2.1 Offline Evaluation in Recommendation

Online A/B tests are effective in tech companies to assess the performances of recommender models [7]. However, it usually consumes much time and money, which makes it impractical for academic researchers to conduct the evaluation online.

Therefore, researchers usually resort to offline evaluation, which includes: (1) computing traditional indicators (such as Precision, Recall, and NDCG [9]) on the test set; (2) conducting *off-policy evaluation* or *counterfactual reasoning* [1, 23] to perform an offline

---

[1]https://www.kuaishou.com/cn

A/B test [7]; (3) using user simulation techniques [8, 30] to evaluate the recommender in the online A/B test.

However, computing traditional indicators cannot answer how the recommender performs on the massive missing data, i.e., the user-item pairs which have not occurred in the test set. Though off-policy evaluation and user simulation can solve this pitfall, they suffer the problem of high-variance issue [19] and additionally introduced error [5], respectively.

2.2 Datasets with Randomly Exposed Items

A way to fundamentally solve this problem in offline evaluation is to collect unbiased data, i.e., to elicit user preferences on the randomly exposed items. We briefly introduce classic existing high-quality datasets that contain unbiased randomly-sampled data used for offline evaluation in recommendation.

- **Yahoo!R3** [16]. It contains the conventional missing-not-at-random (MNAR) data from 15, 400 users on 1, 000 items in total. It also contains a set of missing-complete-at-random (MCAR) data by asking 5, 400 users to give ratings on 10 items that are randomly selected from the 1,000 items. However, it does not contain timestamps or the features of users and items.
- **Yahoo!R6** [11]. It has two versions, namely R6A and R6B, both of which have only the random policy to expose hundreds of news millions of times. Besides, all user IDs are removed, which forbids it to perform basic collaborative filtering tasks in recommendation.
- **Coat** [22]. It collects the ratings of 290 users on 24 self-selected items and 16 randomly-selected items from a total of 300 items. Similar to Yahoo!R3, the scale of it is too small compared with other datasets in real-world recommendation scenarios. Neither does it has the timestamp to support research on sequential recommendation.
- **Open Bandit Dataset** [19]. It contains interactions collected from two logged bandit policies: a Bernoulli Thompson Sampling policy and a uniform (random) policy. There are approximately 26 million interactions collected by users' clicks on 80 items. The disadvantage of this data is the lack of user ID, making it hard to support the research on sequential recommendation. In addition, the fact that it has only 80 items limits its reliability in evaluating most recommendation algorithms.

The statistics of these datasets are summarized in Table 1. Compared with them, our proposed KuaiRand has the following key advantages: (1) it has the most comprehensive side information including explicit user IDs, interaction timestamps, and rich features for users and items; (2) it has 15 policies with each catered for a special recommendation scenario in Kuaishou App; (3) we introduce 12 feedback signals for each interaction to describe the user's comprehensive feedback; (4) each user has thousands of historical interactions on average. These distinctive characteristics can enable a lot of potential research directions.

3 DATA DESCRIPTION

Here, we introduce the basic characteristics of Kuaishou App and the process of how we collect the KuaiRand dataset.
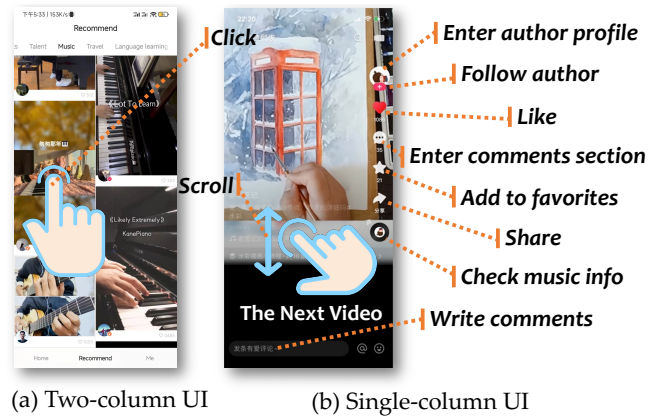

(a) Two-column UI   (b) Single-column UI

Figure 2: Illustration of two recommendation scenarios in Kuaishou App. A variety of user behaviors are marked.

3.1 Characteristics of Kuaishou App

Kuaishou App is one of the largest short video-sharing platforms in China. It has amassed over 300 million daily-active users with an average spent time of nearly 128.1 minutes per day [24]. Different from other recommendation scenarios such as e-commerce, users in short video recommendation platforms can generate very long interaction sequences every day. There are also many new videos published every day, hence the videos become outdated very quickly.

The videos are mainly recommended to users in different 15 scenarios, e.g., under different user interfaces (UIs) in Kuaishou App or the embedded UI in other Apps such as WeChat Mini Program. Each scenario has distinctive rules and objectives. Therefore, each scenario has its own recommendation policy and there are 15 policies in total. We illustrate the most two commonly used UIs in two scenarios in Fig. 2.

On the left-hand side lies the main UI, where the thumbnails of recommended videos are organized in a two-column cascade style. In this UI, users can scroll up or down until they find the video they want to view, and then view this video by clicking the thumbnail. Therefore, *click* is a critical signal indicating user interest in this UI style. On the right-hand side, we can see the single-column recommendation cascading UI. It is the most commonly used UI. The video will automatically play whenever switching to this UI. There is no click signal in the two-column UI, but users can choose to stop viewing this video by exiting this UI or scrolling to the last/next video. Therefore, *view time* in this UI becomes an important signal to reflect user satisfaction. Besides, there are other signals representing users' behaviors when they interact with this video, including *like*, *enter author profile*, *follow author*. These feedback signals correspond to the buttons on this UI (Fig. 2), so we can easily collect these timely signals.

3.2 Data Collection

Our goal is to collect the sequential recommendation data with randomly exposed videos inserted. We achieve this goal via the following three steps:
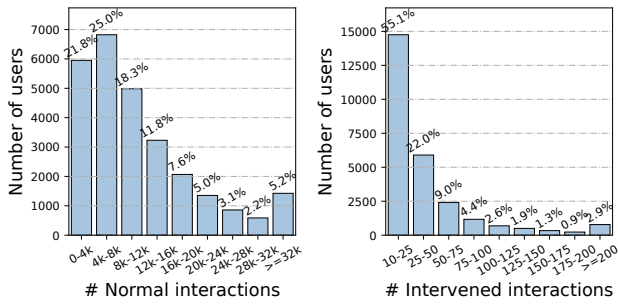
Figure 3: Distribution of users w.r.t. the number of normal interactions (i.e., receiving ordinary recommendations) and intervened interactions (i.e., exposed to random videos).

(i) We sample a batch of videos and filter out the spam such as advertisements. There are $7,583$ items in total. For the target users, we randomly select a batch of users and remove robots, which includes over $200,000$ real users.

(ii) Each time when the recommender system recommends a video list to a user, we decide whether to insert a random item with a fixed probability[2]. If the answer is yes, then we intervene in the recommendation list by randomly selecting one video from this list and replacing it with a random item uniformly sampled from the $7,583$ items. This process lasts two weeks (from April 22, 2022, to May 08, 2022). We record 12 feedback signals reflecting the comprehensive behaviors of these users. We remove the users that have been exposed to less than 10 randomly exposed videos for faithful evaluation. There are $27,285$ users retained. All $7,583$ items have been inserted at least once, and the total number of random interventions is $1,186,059$.

(iii) To facilitate model learning on this dataset, we further retrieve the historical user interaction of these $27,285$ users (without intervention by the random policy), which also lasts two weeks (from April 08, 2022, to April 21, 2022). There are $32,038,725$ items and $322,278,385$ times of normal recommendations included in the one-month dataset. In addition, we further collect the side information including the key features of these users and items.

So far, we have collected the KuaiRand dataset.

3.3 Statistics and Usage

The basic statistics of KuaiRand are summarized in Table 1. We will release three versions (*KuaiRand-27K*, *KuaiRand-1K*, and *KuaiRand-Pure*) for different usages.

For *KuaiRand-27K*, we visualize the distribution of the users w.r.t. the numbers of interactions Fig. 3. An interaction can either be a normally recommended video or a randomly chosen video. From the figure, we can see that most users in this dataset have viewed thousands of recommended videos in one month. Each user has more than 10 times viewed random videos. For the detailed usage and news of this dataset, please refer to https://kuairand.com.

---

[2]We hide the exact value for a commercial security reason, i.e., to avoid leaking out the key parameters of the online recommendation strategy.

4 POTENTIAL RESEARCH DIRECTIONS

Through releasing KuaiRand, we offer the opportunity to carry out the debiasing task in the large-scale sequential recommendation for the first time.

- **Debiasing in Recommendation**. Recommender systems suffer from various biases in the data collection stage [3]. Most existing datasets are very sparse and affected by user-selection bias [16] or exposure bias (previous model bias) [13]. It is of critical importance to develop models that can alleviate biases. To evaluate the models, we need reliable unbiased data. KuaiRand is the first dataset that inserts the random items into the normal recommendation feeds with rich side information and all item/user IDs provided. With this authentic unbiased data, we can evaluate and thus improve the recommender policy.

With its distinctive features, KuaiRand can further support the following promising research directions in recommendation.

- **Off-policy Evaluation (OPE).** OPE is proposed to overcome the difficulties in conducting online A/B testing in recommendation. OPE aims to estimate the performance of a counterfactual policy using only log data generated by past policies [19]. It requires that we have the log sequences rather than merely positive samples. KuaiRand satisfies this requirement. It recorded logs under 15 policies catered for different situations. Therefore, we can leverage OPE and conduct offline A/B tests to improve our recommendation policies [7].

- **Interactive Recommendation.** An interactive recommender system (IRS) is usually formulated as a decision-making process to pursue long-term success in recommendation. IRS built on reinforcement learning has shown superiority to the traditional supervised frameworks [4, 26, 27]. KuaiRand contains the sequential logs generated from different policies that can naturally support the research on IRSs with the help of OPE and user simulation techniques [5, 10].

- **Long Sequential Behavior Modeling.** In practice, traditional sequential recommendation models, such as DIN [32] and DIEN [31], can only model a maximum of hundreds of items in a sequence. Recently, some works have demonstrated that modeling users' long sequential behaviors is very effective in lifting click-through rate (CTR) [17, 18]. KuaiRand has thousands of interactions for each user, which is perfect for the long sequential user modeling task.

- **Multi-Task Learning.** In real-world recommender systems, user satisfaction can be reflected in multiple behaviors, e.g., clicks, writing down positive comments, and staying longer. Usually, these metrics cannot be pursued at the same time in one model. Therefore, multi-task learning or multi-object learning has received growing attention [2, 25, 28]. KuaiRand recorded 12 feedback signals for each interaction, which promotes the research on learning with more than ten tasks. In contrast, most existing datasets in multi-task recommendation [12, 15] only provide less than 5 signals.

5 ACKNOWLEDGEMENTS

REFERENCES

[1] Léon Bottou, Jonas Peters, Joaquin Quiñonero-Candela, Denis X Charles, D Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. 2013. Counterfactual Reasoning and Learning Systems: The Example of Computational Advertising. *Journal of Machine Learning Research* 14, 11 (2013).

[2] Qingpeng Cai, Ruohan Zhan, Chi Zhang, Jie Zheng, Guangwei Ding, Pinghua Gong, Dong Zheng, and Peng Jiang. 2022. Constrained Reinforcement Learning for Short Video Recommendation. *arXiv preprint arXiv:2205.13248* (2022).

[3] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. 2020. Bias and Debias in Recommender System: A Survey and Future Directions. *arXiv preprint arXiv:2010.03240* (2020).

[4] Chongming Gao, Wenqiang Lei, Jiawei Chen, Shiqi Wang, Xiangnan He, Shijun Li, Biao Li, Yuan Zhang, and Peng Jiang. 2022. CIRS: Bursting Filter Bubbles by Counterfactual Interactive Recommender System. *arXiv preprint arXiv:2204.01266* (2022).

[5] Chongming Gao, Wenqiang Lei, Xiangnan He, Maarten de Rijke, and Tat-Seng Chua. 2021. Advances and Challenges in Conversational Recommender Systems: A Survey. *AI Open* 2 (2021), 100–126. https://doi.org/10.1016/j.aiopen.2021.06.002

[6] Chongming Gao, Shijun Li, Wenqiang Lei, Jiawei Chen, Biao Li, Peng Jiang, Xiangnan He, Jiaxin Mao, and Tat-Seng Chua. 2022. KuaiRec: A Fully-observed Dataset and Insights for Evaluating Recommender Systems. *arXiv preprint arXiv:2202.10842* (2022).

[7] Alexandre Gilotte, Clément Calauzènes, Thomas Nedelec, Alexandre Abraham, and Simon Dollé. 2018. Offline A/B Testing for Recommender Systems. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (Marina Del Rey, CA, USA). Association for Computing Machinery, New York, NY, USA, 198–206. https://doi.org/10.1145/3159652.3159687

[8] Eugene Ie, Chih-wei Hsu, Martin Mladenov, Vihan Jain, Sanmit Narvekar, Jing Wang, Rui Wu, and Craig Boutilier. 2019. Recsim: A Configurable Simulation Platform for Recommender Systems. *arXiv preprint arXiv:1909.04847* (2019).

[9] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated Gain-Based Evaluation of IR Techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (Oct. 2002), 422–446. https://doi.org/10.1145/582415.582418

[10] Wenqiang Lei, Chongming Gao, and Maarten de Rijke. 2021. RecSys 2021 Tutorial on Conversational Recommendation: Formulation, Methods, and Evaluation. In *Fifteenth ACM Conference on Recommender Systems* (Amsterdam, Netherlands) *(RecSys '21)*. Association for Computing Machinery, New York, NY, USA, 842–844. https://doi.org/10.1145/3460231.3473325

[11] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. 2010. A Contextual-Bandit Approach to Personalized News Article Recommendation. In *Proceedings of the 19th International Conference on World Wide Web* (Raleigh, North Carolina, USA). Association for Computing Machinery, New York, NY, USA, 661–670. https://doi.org/10.1145/1772690.1772758

[12] Xiao Lin, Hongjie Chen, Changhua Pei, Fei Sun, Xuanji Xiao, Hanxiao Sun, Yongfeng Zhang, Wenwu Ou, and Peng Jiang. 2019. A Pareto-efficient Algorithm for Multiple Objective Optimization in E-commerce Recommendation. In *Proceedings of the 13th ACM Conference on recommender systems (RecSys '19)*. 20–28.

[13] Dugang Liu, Pengxiang Cheng, Zhenhua Dong, Xiuqiang He, Weike Pan, and Zhong Ming. 2020. A General Knowledge Distillation Framework for Counterfactual Recommendation via Uniform Data. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, China). Association for Computing Machinery, New York, NY, USA, 831–840. https://doi.org/10.1145/3397271.3401083

[14] David C. Liu, Stephanie 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) *(WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 583–592. https://doi.org/10.1145/3041021.3054202

[15] 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 '18)*. 1137–1140.

[16] Benjamin M. Marlin and Richard S. Zemel. 2009. Collaborative Prediction and Ranking with Non-Random Missing Data. In *Proceedings of the Third ACM Conference on Recommender Systems* (New York, New York, USA). Association for Computing Machinery, New York, NY, USA, 5–12. https://doi.org/10.1145/1639714.1639717

[17] Qi Pi, Weijie Bian, Guorui Zhou, Xiaoqiang Zhu, and Kun Gai. 2019. Practice on Long Sequential User Behavior Modeling for Click-Through Rate Prediction. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA) *(KDD '19)*. Association for Computing Machinery, New York, NY, USA, 2671–2679. https://doi.org/10.1145/3292500.3330666

[18] Qi Pi, Guorui Zhou, Yujing Zhang, Zhe Wang, Lejian Ren, Ying Fan, Xiaoqiang Zhu, and Kun Gai. 2020. Search-Based User Interest Modeling with Lifelong Sequential Behavior Data for Click-Through Rate Prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (Virtual Event, Ireland) *(CIKM '20)*. Association for Computing Machinery, New York, NY, USA, 2685–2692. https://doi.org/10.1145/3340531.3412744

[19] Yuta Saito, Shunsuke Aihara, Megumi Matsutani, and Yusuke Narita. 2021. Open Bandit Dataset and Pipeline: Towards Realistic and Reproducible Off-Policy Evaluation. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. https://openreview.net/forum?id=tyn3MYS_uDT

[20] Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. 2020. Unbiased Recommender Learning from Missing-Not-At-Random Implicit Feedback. In *Proceedings of the 13th International Conference on Web Search and Data Mining* (Houston, TX, USA). Association for Computing Machinery, New York, NY, USA, 501–509. https://doi.org/10.1145/3336191.3371783

[21] Tobias Schnabel, Paul N. Bennett, Susan T. Dumais, and Thorsten Joachims. 2018. Short-Term Satisfaction and Long-Term Coverage: Understanding How Users Tolerate Algorithmic Exploration. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (Marina Del Rey, CA, USA). Association for Computing Machinery, New York, NY, USA, 513–521. https://doi.org/10.1145/3159652.3159700

[22] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning* (New York, NY, USA). JMLR.org, 1670–1679.

[23] Adith Swaminathan and Thorsten Joachims. 2015. Counterfactual Risk Minimization: Learning from Logged Bandit Feedback. In *International Conference on Machine Learning*. PMLR, 814–823.

[24] CIW TEAM. 2022. Short video app Kuaishou e-commerce GMV up 482022. https://www.chinainternetwatch.com/31784/kuaishou-quarterly/. Accessed June 1, 2022.

[25] Hongwei Wang, Fuzheng Zhang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2019. Multi-Task Feature Learning for Knowledge Graph Enhanced Recommendation. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 2000–2010. https://doi.org/10.1145/3308558.3313411

[26] Shiqi Wang, Chongming Gao, Min Gao, Junliang Yu, Zongwei Wang, and Hongzhi Yin. 2022. Who Are the Best Adopters? User Selection Model for Free Trial Item Promotion. *arXiv preprint arXiv:2202.09508* (2022).

[27] Xin Xin, Alexandros Karatzoglou, Ioannis Arapakis, and Joemon M. Jose. 2020. Self-Supervised Reinforcement Learning for Recommender Systems. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery, New York, NY, USA, 931–940. https://doi.org/10.1145/3397271.3401147

[28] Chenxiao Yang, Junwei Pan, Xiaofeng Gao, Tingyu Jiang, Dapeng Liu, and Guihai Chen. 2022. Cross-Task Knowledge Distillation in Multi-Task Recommendation. *AAAI '22* (2022).

[29] Longqi Yang, Yin Cui, Yuan Xuan, Chenyang Wang, Serge Belongie, and Deborah Estrin. 2018. Unbiased Offline Recommender Evaluation for Missing-Not-at-Random Implicit Feedback. In *Proceedings of the 12th ACM Conference on Recommender Systems* (Vancouver, British Columbia, Canada). Association for Computing Machinery, New York, NY, USA, 279–287. https://doi.org/10.1145/3240323.3240355

[30] Shuo Zhang and Krisztian Balog. 2020. Evaluating Conversational Recommender Systems via User Simulation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, NY, USA, 1512–1520. https://doi.org/10.1145/3394486.3403202

[31] Guorui Zhou, Na Mou, Ying Fan, Qi Pi, Weijie Bian, Chang Zhou, Xiaoqiang Zhu, and Kun Gai. 2019. Deep Interest Evolution Network for Click-through Rate Prediction. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence* (Honolulu, Hawaii, USA) *(AAAI '19)*. AAAI Press, Article 729, 8 pages. https://doi.org/10.1609/aaai.v33i01.33015941

[32] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep Interest Network for Click-Through Rate Prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (London, United Kingdom). Association for Computing Machinery, New York, NY, USA, 1059–1068. https://doi.org/10.1145/3219819.3219823