Challenges for Session-based Recommender Systems in next generation IFE-Systems

Marko Harasic
marko.harasic@fokus.fraunhofer.de
Fraunhofer FOKUS

Adrian Paschke
adrian.paschke@fokus.fraunhofer.de
Fraunhofer FOKUS

ABSTRACT

This position paper describes the arising challenges and requirements for providing recommendations to travelers in airplanes. Recommender systems are an integral part of modern-day user experience when users choose between plenty of possible items. They understand their preferences and support them in discovering meaningful content by creating personalized recommendations. The next generation of In-flight Entertainment systems will provide travelers with the choice of several thousands multimedia items. Given the constraints of that setting, traditional recommender systems can not run on board an airplane. This paper identifies the challenges and provides the first step in research to realize a system in such an environment. Session-based recommender systems detect the users' intents quickly and are well suited to the challenge of anonymous users. Furthermore, they utilize the information from all users and provide them Wisdom of the Crowd-based recommendations. Unfortunately, their required computational resources and information flow require a connection between the participants and the central instance. Such a connection does not exist in airplanes because of security reasons and added weight. That creates a significant challenge to provide recommendations to the users if only local devices are available without a stable network. Federated Learning is a modern approach for training, updating, and operating novel Machine Learning approaches in distributed environments. The combination of modern Recommender Systems and Federated Learning is a promising approach to realize the a system for next-generation IFE-systems. Special evaluation methods need to be developed, and the first evaluation concepts are presented.

KEYWORDS

Recommender Systems, Session-based Recommender Systems, Federated Learning, Edge Computing, Evaluation Protocols

ACM Reference Format:
Marko Harasic and Adrian Paschke. 2022. Challenges for Session-based Recommender Systems in next generation IFE-Systems. In *Proceedings of The 2nd International Workshop on Online and Adaptive Recommender Systems (⟨KDD '22 OARS Workshop⟩)*. ACM, New York, NY, USA, 5 pages.

⟨KDD '22 OARS Workshop⟩, ⟨14th August 2022⟩, ⟨Washington D.C., US⟩

1 INTRODUCTION

Recommender Systems (RSs) are today an integral part of modern-day information portals, streaming websites, and e-commerce solutions. Traced back in the field of Information Retrieval (IR) [32], users inform themselves, book holidays, consume multimedia content, or shop items. E-commerce websites like Amazon list over 12 million products that the user can choose from, and even the 14.000 movies and series on Netflix create a decision challenge for the customer. RSs provide a mechanism to support the users in their task of navigation and decision making by proposing recommendations on items that fit their preferences and needs. They analyze the user's behavior and generate profiles, which map their interests to an internal representation called the user profile. Furthermore, the personalized recommendations created by RSs provide a significant business benefit by driving sales and introducing new content for the consumers. Approximately 30% of Amazons' view come from recommendations [36] and more than 80% of Netflix watching events are created by personalized recommendations [12].

Unfortunately, those algorithms run in large data centers and need a stable network connection to capture the users' interactions to create a meaningful profile. But what if no or only a weak network is available? Today's modern personal devices can store thousands of entries such as music or movies. Users do not want to miss the luxury of personalization on their devices when they create a playlist of songs for a specific need or watch certain content based on their current mood. Even though IR techniques can be applied in that kind of setting, Wisdom of the Crowd [38] approaches have significant benefits because of those tasks' highly subjective nature. A case study for further research was defined in collaboration with an industrial partner to provide a real-world setting. This paper describes the first insights and challenges to address.

2 CASE STUDY: IFE-RECOMMENDER

Millions of passengers take a flight each day, and their entertainment is a key aspect of modern-day air travel. It was shown that a positive multimedia experience is beneficial for the stress level of each passenger in his travels. From the first in-flight movies screened in 1921 at an exposition in Chicago to single-aisle coach-style television systems in the late 1970s, today's In-Flight Entertainment (IFE) systems offer each passenger their own personal system. Today, multimedia content in airplanes is usually limited. Frequent travelers consume all available content in a short time, and after a few successive travels, there is no unseen content available for them. However, the deployment of next-generation IFE systems with thousands of movies available for travelers is the obvious step forward.

Recently, providers of IFE systems started to extend the availability of multimedia content in the airplanes and started to host several hundred to thousands of items. Given the fact that travelers now have access to that large corpus of content, they face the same challenges as browsing on streaming websites like Netflix with their hosted 14.000 movies and series. Travelers are confronted with plenty of choices, overwhelmed by the possibilities, and face the Paradox of Choice [35] and having only the small devices installed in the seats in front of them. Therefore, they need nowadays support from appropriate tools to access and navigate through the vast space of items [39]. RSs are the natural choice to provide users an improved experience by creating recommendations for items that match their personal preferences.

2.1 Problem setting for IFE-Recommenders

During the period of a flight, only a few interactions happen between the traveler and the system. Usually, more time is spent consuming content than browsing, making it challenging to gather enough data to build a comprehensive profile. Furthermore, travelers are anonymous, and no persistent storage of derived preferences is created.

Recommendations can not be calculated by a central compute server in the airplane, and the interaction data from the local device can not be transmitted to the server. IFE-devices in planes usually do not have network connectivity during flight time. Therefore, local devices have to generate these recommendations for the traveler. Those – mostly ARM-based – devices have limited resources in available memory and computational capabilities and can not operate on the typical large models, which are today applied by providers of personalized services.

Even though these devices could operate on local and streamlined models, updates regarding new information can not be deployed without addressing further challenges. Airplanes usually communicate over satellite, and connections with higher bandwidth to upload all user interactions and update the models are seldomly accessible. With a stable connection, the devices could upload the interactional data to a central training server, the model retrained, and the new model could be rolled out. Unfortunately, the training of modern-day Deep Learning (DL)-models is resource-intensive, and there exists a high probability that the new model is not deployed in the timeslot when the aircraft has a stable connection to the central instance. Therefore, a delay between the insights taken from the recent interactional data and the deployment of the updated model. An old model from a former upload-retrain-deploy cycle would then calculate recommendations with lower quality.

Furthermore, IFE system manufacturers, as well as multimedia, capture the user action such as browsing in the item base, interaction patterns, and watching behavior. Such customer data is highly precious, giving those companies insights and support for strategic decisions [8]. The value of those companies directly correlates to the worth of their user base, making them with obvious reasons not willing to share it [13]. At the same time, a vast userbase can benefit all collaborators, and everyone could profit from the competitors as well.

2.2 Challenges and possible approaches for IFE-Recommenders

In the following, the questions which were derived in the planning stage of the IFE-Recommender are presented. Furthermore, recent research's possible solutions and best practices to address the challenges are described.

2.2.1 How anonymous users can get good recommendations? According to their kind of calculation approach, RSs are classified into content based filtering (CBF) and into collaborative filtering (CF) [1]. In CBF, that type of RS determines items that have similar features to those items preferred in the past. CF follows the assumption that similar users have similar preferences. Then the system proposes to the current user possibly interesting items which are liked by those peers. CF has the advantage of considering only the users' point of view, and that method doesn't need to extract the features of the items to match them against the users' preferences. Therefore, they are highly domain-independent and can be used to recommend items independently of their type [1]. Unfortunately, those classical approaches rely on static representations of user preferences in the form of their profiles. Those profiles are built by collecting past data and therefore make it challenging to operate such a RS in a setting where the interactions are ephemeral and anonymous.

Session-Based Recommender Systems (SBRSs) are designed to discover the intent of anonymous users in their interaction with the system after a short time with the need of a limited amount of data points. They take only short interaction sequences (so-called session) of usually 4 to 10 events into consideration to derive the interactional view [10] of the current users' context. That kind of RS treats each session independently, and without the need for a persistent user profile [41]. As no profiles are needed, SBRSs are a powerful tool to address the challenge of anonymous or new users [41]. That isolated treatment of sessions allows discovering the user's goal independently. As SBRSs address these challenges by design, their deployment is the natural choice to realize a next-generation IFE system.

2.2.2 Which is the most suitable IFE-Recommender SBRS approach? SBRSs are classified into traditional heuristic-based approaches, into factorization-based methods and into DL-based methods [41]. Factorisation Machines (FMs) [31] are an extension of Matrix Factorisation (MF) [17], which were originally developed to particularly address the problem of sparse data. SBRSs consider the user ratings and further contextual information of the users' interaction with the system as an input to create recommendations. With the recent advances in DL [20] in fields such as in computer vision [18] or natural language processing [34], it was a matter of time until DL-algorithms were deployed in the field of RS. Especially when it comes to analyse and interpret sequential data such as the interaction patterns occurring in SBRSs, DL methods based on Recurrent Neural Network (RNN) – a neural network architecture, which covers the sequential behavior of data – could improve the accuracy of recommendations [15]. Being successfully applied in NLP-settings and the similar sequential structure of sessions and sentences, Transformers such as BERT4Rec [37] and STAMP [22] were recently adapted to the recommendation task in SBRS.

DL based methods got increased popularity in the field of RSs, outperforming traditional approaches in several applications [43]. Controversially, the opposite was found in a recent evaluation, as MF approaches still can exceed the more modern DL algorithms in the investigated domains [7]. While those evaluations happened in the setting of traditional RS, it should be investigated which of those two results holds in the context of SBRSs in the IFE-setting.

2.2.3 Can Wisdom of the Crowd be used in a distributed environment without stable network connections? In the traditional centralized approach for DL, data collected by devices is uploaded and processed centrally on a cloud-based server or in a data center. The model is updated with local information (images, location information, or sensor values) and then distributed among the participants. Originally motivated by privacy challenges and the availability of computational resources in modern-day personal devices, a decentralized machine learning approach called Federated Learning (FL) was developed by Google [24]. By its architecture, FL guarantees that training data remains on personal devices while the users collaboratively train a shared model. The devices send only model updates (weights of the model or gradients of a feed-forward model from the loss function) to a central instance. That system aggregates the particular results and transfers the updates back to the participants. This step is repeated until a desirable precision is achieved. Then, each device has the refined model locally stored and performs further calculations independently of a central instance or network availability.

The research of FL in the context of RSs is an emerging new field [21]. Personalization happens directly at the device, and only the interpretations of actions are shared between the peers. With their locally trained profile but contributing to the global model in the form of updates with small network requirements, peers do not need constantly upload their preferences but still benefit from the Wisdom of the Crowd. As MF is still considered state of the art, it is obvious that the first research in the combination of RS and FL bases on MF approaches [4]. Nevertheless, the first implementations of the federated DL approach and their evaluation comparing centralized algorithms were published recently. The surveys of Yang et al. [42] and Alamgir et al. [3] give a comprehensive overview on current trends and open research questions in the field of Federated RS.

2.2.4 How to share the information between different peers? Besides the ability to update models asynchronously and to generate the recommendations locally, FL allows sharing the insights of the models between different peers, such as the provider of IFE-solutions and the content providers. FL for RS are distinguished into vertical, horizontal, and transfer FL [42]. To address this challenge, horizontal FL is the most suitable approach when all peers have data coming from the same domain and share common identifiers for the recommendable items. Each party has its particular collection of interactions and trains its model independently without information disclosure of its business data and a common userbase.

2.2.5 How to create a proper baseline? The necessity of "real" user data because of the drawbacks of synthetic datasets is a well-known problem [2]. With the growing corporate interest in SBRS, several data sets in different domains capturing sequential browsing and consuming behavior of the users were released recently. Unfortunately, researchers focus today only on presenting improved results of their new algorithms compared to older approaches and omit to tune the baselines to their optimal results [6]. A first step on the comprehensive evaluation of SBRSs was performed by Ludewig [23]. Following the evaluation guidelines from that survey, new approaches should be evaluated with data from different published datasets based on real user behavior. Each dataset was derived from different domains and has its own hidden structure of interactions. As the algorithms perform differently on each dataset, they have to be truthfully tuned accordingly to reach their best performance. Only that approach makes an objective analysis of the advantages of new algorithms possible. To span different domains, it is planned to use the following data sets from e-commerce and multimedia together with a dataset created by the industrial partner of traditional IFE-systems to compare the next generation IFE-system objectively:

- **RSC15**, a dataset containing click sequences and shopping actions, which was published in the RecSys 2015 challenge and already been used in the work of Hidasi [15]
- **#nowplaying-RS** was created in order to allow an objective benchmark for evaluating contextual-aware music RS [30]
- **Booking.com trips**, a dataset created for the WSDM 2021 challenge by Booking.com consisting of 359k journeys over 39k destinations [11]
- **Fly-data** A dataset containing traveler sessions with watching events (Play-Stop-Resume) for multimedia content of airplane travelers, which was created by a private company. That dataset will be the input base for the planned IFE-Recommender.

2.2.6 Is accuracy the only critical metric? Even though an offline evaluation has its particular drawbacks as it relies on heuristics with assumptions, that step gives valuable insights before the roll-out of RS systems. Well regarded metrics that measure the accuracy of the predicted content such as Precision@N [14] and ranking metrics such as NDCG@N [40] should be used. Furthermore, systems should be evaluated beyond accuracy as it is nowadays commonly believed that the diversity of recommendations has an even higher importance [25]. By emphasizing the systems' accuracy only, a system would be realized, which generates only boring and ineffective recommendations. That will lead to a limited scope filter bubble which traps the user and hinders his personal growth [28]. Therefore, novelty and the closely related concept of serendipity are central evaluation dimensions when it comes to the quality of the IFE-Recommender [9, 16].

2.2.7 The long-tail and non-IID data. Closely related to the question of serendipity is the behavior of the systems in the so-called long tail of items. The long tail describes the shape of item distribution where only a few items have high popularity (f.e. ratings, sales, or accesses), and most of them are rarely accessed [5]. First formulated by Park and Tuzhilin as the long tail-problem [29], most RS algorithms lose their accuracy when it comes to the access of items in the tail. Such systems behave comparable to simple popularity-based RS. Therefore, performance evaluations have to regard the items' distribution to cover unpopular items. Furthermore, as FL seriously suffers under the non-IID challenge typically arises because of the long-tail data distribution in the RS setting [42], methods to leverage that impact has to be incorporated into the system[44].

2.2.8 How to determine the minimal bound of required computational resources? As the system's performance is significant when it comes to real-world operations, its behavior in dependency on the total amount of users and items has to be investigated according to response time, and throughput [33]. Most modern personal devices, as well as the hardware of IFE-systems, are based on the ARM-architecture. Only a few research concerning the deployment and operation of DL-algorithms have been performed on Edge-Devices [27] and those examinations mainly focused on computer vision [19], and no evaluation of SBRSs on that hardware was performed to the authors best knowledge. A heterogeneous cluster of ARM-based devices with different hardware specifications could be created. The models will be deployed on those ARM-based devices to determine their performance concerning the specific hardware environments. This environment will measure the impact of the model sizes and algorithms on real-world hardware. Those operational parameters are used to simulate a network of FL-nodes.

As the number of nodes in a FL-based RS easily surpasses hundreds of nodes, an evaluation running on real devices on such a scale is not feasible. Therefore, that network of devices has to be simulated. A central instance such as a virtual machine (VM) will provide the environment with several simulated clients reflecting the hardware specifications of the clusters peers to give a good estimation for a real-world distribution. Overhead costs such as network traffic will be simulated and taken into consideration. The systems should be evaluated with the formerly mentioned real-world data sets according to several quantitative metrics covering the accuracy and other metrics beyond the sole accuracy. Furthermore, the correlation between the resource requirements of the investigated algorithms and their overall applicability to the data sets' domain has to be determined. That information brings crucial insights into recommendation quality vs. needed resources and can help to find optimal reduced model sizes – for example, with model pruning [26] – given the hardware constraints.

3 SUMMARY

This work describes the first steps and research questions to realize a case study for a distributed SBRS in an edge environment. That IFE-recommender gives domain-dependent requirements for its operation, such as independence of a stable network, certain hardware constraints, and the existence of anonymous user interactions. This case study can easily be adapted to other domains, such as a playlist generation for local song collections or touristic tour guidance in offline environments. This paper contains an overview of challenges to be addressed in designing such a system, and recently developed methods to solve them are presented. It provides an interesting overview of several state-of-the-art approaches for further research.

4 ACKNOWLEDGMENT

Research for this paper was partially funded by the German Aerospace Center (DLR) project CANARIA under project number 20D1931E. he project partners Safran passenger Innovations, TU Braunschweig, TU Dresden, Cadami GmbH and Fraunhofer FOKUS work together in the CANARIA joint project. We thank our project partners for their contributions and their collaboration to this research work.

REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING* 17, 6 (2005), 734–749.

[2] Charu C. Aggarwal, Joel L. Wolf, Kun-Lung Wu, and Philip S. Yu. 1999. Horting Hatches an Egg: A New Graph-theoretic Approach to Collaborative Filtering. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '99)*. ACM, New York, NY, USA, 201–212. https://doi.org/10.1145/312129.312230

[3] Zareen Alamgir, Farwa K. Khan, and Saira Karim. 2022. Federated recommenders: methods, challenges and future. *Cluster Computing* (June 2022). https://doi.org/10.1007/s10586-022-03644-w

[4] Muhammad Ammad-ud din, Elena Ivannikova, Suleiman A. Khan, Were Oyomno, Qiang Fu, Kuan Eeik Tan, and Adrian Flanagan. 2019. Federated Collaborative Filtering for Privacy-Preserving Personalized Recommendation System. *CoRR* abs/1901.09888 (2019). https://doi.org/10.48550/arXiv.1901.09888

[5] Chris Anderson. 2006. *The long tail: why the future of business is selling less of more* (1st ed ed.). Hyperion, New York.

[6] Vito Walter Anelli, Alejandro Bellogín, Tommaso Di Noia, Dietmar Jannach, and Claudio Pomo. 2022. Top-N Recommendation Algorithms: A Quest for the State-of-the-Art. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization* (Barcelona, Spain) *(UMAP '22)*. Association for Computing Machinery, New York, NY, USA, 121–131. https://doi.org/10.1145/3503252.3531292

[7] Vito Walter Anelli, Alejandro Bellogín, Tommaso Di Noia, and Claudio Pomo. 2021. Reenvisioning the comparison between Neural Collaborative Filtering and Matrix Factorization. In *Fifteenth ACM Conference on Recommender Systems*. Association for Computing Machinery, New York, NY, USA, 521–529. https://doi.org/10.1145/3460231.3475944

[8] Josh Brownlow, Mohamed Zaki, Andy Neely, and Florian Urmetzer. 2015. Data-Driven Business Models: A Blueprint for Innovation.

[9] Pablo Castells, Neil J. Hurley, and Saul Vargas. 2015. Novelty and Diversity in Recommender Systems. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer US, Boston, MA, 881–918. http://link.springer.com/10.1007/978-1-4899-7637-6_26

[10] Paul Dourish. 2004. What we talk about when we talk about context. *Personal and Ubiquitous Computing* 8, 1 (Feb. 2004), 19–30. https://doi.org/10.1007/s00779-003-0253-8

[11] Dmitri Goldenberg and Pavel Levin. 2021. Booking.com Multi-Destination Trips Dataset. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery, New York, NY, USA, 2457–2462. https://doi.org/10.1145/3404835.3463240

[12] Carlos A. Gomez-Uribe and Neil Hunt. 2016. The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Transactions on Management Information Systems* 6, 4 (Dec. 2016), 13:1–13:19. https://doi.org/10.1145/2843948

[13] Andy Green. 2012. Understanding the value of customer data. *Journal of Direct, Data and Digital Marketing Practice* 13, 3 (Jan. 2012), 221–233. https://doi.org/10.1057/dddmp.2011.47

[14] Guibing Guo. 2012. Resolving Data Sparsity and Cold Start in Recommender Systems. In *User Modeling, Adaptation, and Personalization*, Judith Masthoff, Bamshad Mobasher, Michel C. Desmarais, and Roger Nkambou (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 361–364.

[15] Balázs Hidasi and Alexandros Karatzoglou. 2018. Recurrent Neural Networks with Top-k Gains for Session-Based Recommendations. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (Torino, Italy) *(CIKM '18)*. Association for Computing Machinery, New York, NY, USA, 843–852. https://doi.org/10.1145/3269206.3271761

[16] Marius Kaminskas and Derek Bridge. 2017. Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. *ACM Transactions on Interactive Intelligent Systems* 7, 1 (March 2017), 1–42. https://doi.org/10.1145/2926720

[17] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer* 42, 8 (Aug. 2009), 30–37. https://doi.org/10.1109/MC.2009.263

[18] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (2017), 84–90. https://doi.org/10.1145/3065386

[19] Liangzhen Lai, Naveen Suda, and Vikas Chandra. 2018. *CMSIS-NN: Efficient Neural Network Kernels for Arm Cortex-M CPUs*. Technical Report arXiv:1801.06601.

arXiv. https://doi.org/10.48550/arXiv.1801.06601 arXiv:1801.06601 [cs] type: article.

[20] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep Learning. *Nature* 521, 7553 (2015), 436–444. https://doi.org/10.1038/nature14539

[21] Qinbin Li, Zeyi Wen, Zhaomin Wu, Sixu Hu, Naibo Wang, Yuan Li, Xu Liu, and Bingsheng He. 2021. A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection. *IEEE Transactions on Knowledge and Data Engineering* (2021), 1–1. https://doi.org/10.1109/TKDE.2021.3124599

[22] Qiao Liu, Yifu Zeng, Refuoe Mokhosi, and Haibin Zhang. 2018. STAMP: Short-Term Attention/Memory Priority Model for Session-Based Recommendation. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18)*. Association for Computing Machinery, New York, NY, USA, 1831–1839. https://doi.org/10.1145/3219819.3219950

[23] Malte Ludewig, Noemi Mauro, Sara Latifi, and Dietmar Jannach. 2021. Empirical analysis of session-based recommendation algorithms. *User Modeling and User-Adapted Interaction* 31, 1 (March 2021), 149–181. https://doi.org/10.1007/s11257-020-09277-1

[24] H. Brendan McMahan, Eider Moore, Daniel Ramage, and Blaise Agüera y Arcas. 2016. Federated Learning of Deep Networks using Model Averaging. *CoRR* abs/1602.05629 (2016). http://arxiv.org/abs/1602.05629

[25] Sean M. McNee, John Riedl, and Joseph A. Konstan. 2006. Being Accurate is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI EA '06)*. Association for Computing Machinery, New York, NY, USA, 1097–1101. https://doi.org/10.1145/1125451.1125659

[26] Pavlo Molchanov, Stephen Tyree, Tero Karras, Timo Aila, and Jan Kautz. 2017. *Pruning Convolutional Neural Networks for Resource Efficient Inference*. Technical Report arXiv:1611.06440. arXiv. https://doi.org/10.48550/arXiv.1611.06440 arXiv:1611.06440 [cs, stat] type: article.

[27] M. G. Sarwar Murshed, Christopher Murphy, Daqing Hou, Nazar Khan, Ganesh Ananthanarayanan, and Faraz Hussain. 2022. Machine Learning at the Network Edge: A Survey. *Comput. Surveys* 54, 8 (Nov. 2022), 1–37. https://doi.org/10.1145/3469029 arXiv:1908.00080 [cs, stat].

[28] Eli Pariser. 2011. *The filter bubble: what the Internet is hiding from you*. Penguin Press, New York. OCLC: ocn682892628.

[29] Yoon-Joo Park and Alexander Tuzhilin. 2008. The Long Tail of Recommender Systems and How to Leverage It. In *Proceedings of the 2008 ACM Conference on Recommender Systems (RecSys '08)*. ACM, New York, NY, USA, 11–18. https://doi.org/10.1145/1454008.1454012

[30] Asmita Poddar, Eva Zangerle, and Yi-Hsuan Yang. 2018. #nowplaying-RS: A New Benchmark Dataset for Building Context-Aware Music Recommender Systems. In *Proceedings of the 15th Sound & Music Computing Conference*. Zenodo, Limassol, Cyprus. http://mac.citi.sinica.edu.tw/~yang/pub/poddar18smc.pdf

[31] Steffen Rendle. 2010. Factorization Machines. In *2010 IEEE International Conference on Data Mining*. IEEE Computer Society, Washington, DC, USA, 995–1000. https://doi.org/10.1109/ICDM.2010.127 ISSN: 2374-8486.

[32] Gerard Salton. 1988. *Automatic text processing: the transformation, analysis, and retrieval of information by computer*. Addison-Wesley, Reading, Mass.

[33] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based Collaborative Filtering Recommendation Algorithms. In *Proceedings of the 10th International Conference on World Wide Web (WWW '01)*. ACM, New York, NY, USA, 285–295. https://doi.org/10.1145/371920.372071

[34] Jürgen Schmidhuber. 2015. Deep learning in neural networks: An overview. *Neural Networks* 61 (Jan. 2015), 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

[35] Barry Schwartz. 2007. *The paradox of choice: why more is less* (reissued ed.). Harper Perennial, New York, NY.

[36] Amit Sharma, Jake M. Hofman, and Duncan J. Watts. 2015. Estimating the Causal Impact of Recommendation Systems from Observational Data. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation (EC '15)*. Association for Computing Machinery, New York, NY, USA, 453–470. https://doi.org/10.1145/2764468.2764488

[37] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*. Association for Computing Machinery, New York, NY, USA, 1441–1450. https://doi.org/10.1145/3357384.3357895

[38] James Surowiecki. 2004. *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations* (1st ed ed.). Doubleday, New York.

[39] Alexandra Uitdenbogerd and Ron Schyndel. 2002. A review of factors affecting music recommender success. In *ISMIR 2002, 3rd International Conference on Music Information Retrieval, Proceedings*. IRCAM - Centre Pompidou, Paris, France, 5.

[40] Saúl Vargas and Pablo Castells. 2011. Rank and Relevance in Novelty and Diversity Metrics for Recommender Systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems (RecSys '11)*. ACM, New York, NY, USA, 109–116. https://doi.org/10.1145/2043932.2043955

[41] Shoujin Wang, Longbing Cao, Yan Wang, Quan Z Sheng, Mehmet A Orgun, and Defu Lian. 2021. A survey on session-based recommender systems. *ACM Computing Surveys (CSUR)* 54, 7 (2021), 1–38.

[42] Liu Yang, Ben Tan, Vincent W. Zheng, Kai Chen, and Qiang Yang. 2020. Federated Recommendation Systems. In *Federated Learning: Privacy and Incentive*, Qiang Yang, Lixin Fan, and Han Yu (Eds.). Springer International Publishing, Cham, 225–239. https://doi.org/10.1007/978-3-030-63076-8_16

[43] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. 2019. Deep Learning Based Recommender System: A Survey and New Perspectives. *ACM Comput. Surv.* 52, 1 (Feb. 2019), 38. https://doi.org/10.1145/3285029

[44] Hangyu Zhu, Jinjin Xu, Shiqing Liu, and Yaochu Jin. 2021. *Federated Learning on Non-IID Data: A Survey*. Technical Report arXiv:2106.06843. arXiv. http://arxiv.org/abs/2106.06843 arXiv:2106.06843 [cs] type: article.