



Research directions in session-based and sequential recommendation

A preface to the special issue

Dietmar Jannach¹ · Bamshad Mobasher² · Shlomo Berkovsky³

Published online: 6 August 2020
© Springer Nature B.V. 2020

1 Introduction

Recommender systems are software applications that make tailored item suggestions to users, usually with the goal of helping them overcome information overload or make informed choices. This tailoring process is typically based on the assumption that long-term preference information about the individual users is available to the system, most commonly in the form of a user-item rating matrix. In such a setting, the recommendation problem can be abstracted as a “matrix filling” task (Resnick et al. 1994), and often also the information about the time when the preferences were collected and when the recommendations should be delivered are not considered.

In many real-world recommendation scenarios, however, these assumptions might not hold and therefore represent an abstraction of a limited suitability. On the one hand, there can be a substantial number of first-time visitors or anonymous users requesting a recommendation. Clearly, no long-term preference information is available for such users. In this case, providing a tailored recommendation can only be done based on the interactions observed in the ongoing session. On the other hand, when long-term user preference information is available, the temporal dimension of this information may play a role. More recent interactions, for example, might be more relevant than the older ones. In addition, there might also be some ordering constraints among recommendations, e.g., not recommending a mobile phone

✉ Dietmar Jannach
dietmar.jannach@aau.at

Bamshad Mobasher
mobasher@cs.depaul.edu

Shlomo Berkovsky
shlomo.berkovsky@mq.edu.au

¹ University of Klagenfurt, Klagenfurt, Austria

² DePaul University, Chicago, USA

³ Macquarie University, Sydney, Australia

case before the mobile phone itself is bought; or not recommending a movie sequel before the movie itself is watched.

Given the practical importance of these problems, there has been increased attention to *sequence-aware recommender systems* in recent years (Quadrona et al. 2018). These systems differ from many legacy recommenders both in terms of the input they process and the way they consider sequential information when generating the recommendations. Regarding the input, sequence-aware recommenders assume that the available preference information is chronologically ordered or contains timestamps. As for the outputs, sequence-aware methods take this ordering into account when generating the recommendations, e.g., by focusing on recent data, making repeated recommendations, or identifying temporal dependencies in the data.

Both session-based and sequential recommendation approaches can be seen as special cases of sequence-aware methods.

- The main characteristic of the *session-based* approaches is that the tailoring of the recommendations is focused solely on the ongoing user session and, usually, no longer-term information about the user is available.
- *Sequential* approaches incorporate the ordering of the past events when predicting the next ones. Typical examples include the recommendation of the next point-of-interest (POI), given the POIs the user visited previously.

The distinction between session-based and sequential recommendations is not always clear in the literature. Session-based recommendation is a sequential recommendation problem both in the sense that the goal is to predict the next interactions in a given session and that the available data are time ordered. However, not all sequential approaches are session based. Many of them may be based on time-stamped and longer-term user preference information, e.g., those making next-POI or next-item recommendations (Rendle et al. 2010).

Hence, this special issue opts for a binocular view that focuses on both the session-based and sequential recommender systems. In this preface, we first provide a historical overview of sequence-aware recommendations and discuss the recently emerging trends, and then briefly introduce the accepted papers Symeonidis et al. (2020), Sánchez and Bellogín (2020), Lacić et al. (2020).

2 History, trends and research directions

Session-based recommenders have become a trending topic in the recent years. However, despite the practical value of such recommendations in real-life systems, the area has been relatively under-explored.

2.1 History and trends

In practical environments, session-based recommendations have been very common. Consider, e.g., Amazon's "customers who bought ..." recommendation scenario

solely based on the currently viewed item or Last.fm's music recommendations that create playlists from a single seed track. In the academic literature, early works studied session-based recommenders—although not using this term—as a Web navigation support (Mobasher et al. 2002) and a travel planning decision-making tool (Ricci and Del Missier 2004). In the e-commerce domain, session-based recommendations were addressed later on as a machine learning model based on Markov chains by Shani et al. (2005). Sequential playlist generation were the focus of works by Ragno et al. (2005) and Hariri et al. (2012).

More recently, much of the research has focused on the e-commerce domain, primarily due to the availability of new large datasets released by companies. In 2014, some works based on session logs provided by a European fashion retailer were published (Tavakol and Brefeld 2014; Jannach et al. 2015). In 2015, YOOCHOOSE released a large dataset for of the ACM Conference on Recommender Systems (RecSys) challenge. The teams that participated in the challenge mostly used non-neural machine learning approaches. Soon after, however, Hidasi et al. (2016) proposed the landmark GRU4REC method based on recurrent neural networks. This method and the wider availability of public datasets marked a turning point, leading a large number of deep learning based approaches for session-based recommendations. Many of these used GRU4REC and the evaluation methodology of Hidasi et al. (2016) as a comparison baseline.

2.2 Research directions

The broad adoption and of success of GRU4REC and related neural methods—including methods based on alternative architectures such as attention (Li et al. 2017)—had a strong impact on the research community. These works inspired extensive research on a highly relevant problem and nowadays serve as baselines to benchmark new methods, both neural and non-neural ones (Ludewig et al. 2019). Despite this progress, there are a number of problems that require further research. In the following paragraphs, we outline a selection of possible research directions, which, in our opinion, have not received sufficient attention so far.

Incorporating long-term preference information In the recent years, we observed a strong focus on session-based techniques (i.e., on anonymous sessions), where datasets similar to the one released by YOOCHOOSE serve as a basis for evaluations. In real-world settings, however, additional information about past user sessions is often available, either because the users had logged in or because they were tracked with cookies. The number of works that try to incorporate such longer-term preferences in *session-aware* recommenders (Quadrana et al. 2018) is still scarce. Existing works include Quadrana et al. (2017) and Phuong et al. (2019), but these focus on specific domains, e.g., job or music recommendations. Hence, there is an opportunity to study session-aware recommendations using both on the user's short-term intents and their longer-term profile. Furthermore, knowing past user preferences can also help to implement features like reminders (Jannach et al. 2017) or proactive recommendations (Rook et al. 2020).

Use of item and context side information Session-based recommendation essentially is a cold-start problem, where the recommender neither has access to the user's long-term preferences nor to their current situational context. Considering additional side information about the properties of the recommendable items is a proven approach to deal with such situations and has also been explored in previous research (Tavakol and Brefeld 2014; Hariri et al. 2012). In the context of session-based recommendation, more work on integrating additional data sources (such as Hidasi et al. 2016), and building hybrid and context-aware systems is still needed. In recent years, more and more datasets containing such side information have become available, e.g., through challenges such as the CIKM Cup 2016 that shared a dataset containing various item features (CIKM 2016). The key issue, however, is to develop hybrid and context-aware approaches that generalize beyond a given application domain like e-commerce or news recommendations. A particularly interesting research direction may also target navigational actions like catalog browsing or the search functionality.

User perception and impact analysis Research in sequential and session-based recommendation largely exploits offline evaluation methodologies, aiming to improve prediction accuracy on a held-out part of a historical dataset. This methodology, however, has some inherent limitations. While being able to predict whether an item is relevant or not is important for any recommender, offline metrics cannot inform how the recommendations are perceived by users, e.g., whether they help users in their decision-making or if they facilitate item discovery. Considering broader recommendation tasks, questions of user perception have been explored (Ekstrand et al. 2014) and different evaluation frameworks were proposed (Knijnenburg et al. 2012; Pu et al. 2011). The number of published user studies on sequential and session-based recommendation is, however, very limited. Examples of recent studies include (Ludewig and Jannach 2019) and (Kamehkhosh and Jannach 2017), both in the music domain. More work is required to understand how different quality factors (e.g., diversity or discovery support) influence user perceptions.

Explanations of session-aware recommendations Topics associated with explainability of both session-based and sequential recommendations are of a particular importance in practical systems. The former, typically being based only on the current user-recommender interaction session, are critical for instilling user trust in the system and increasing the chances of the user actually following the recommendations (Berkovsky et al. 2017). Sequential recommenders, on the contrary, incorporate longer-term user profiles and also take into consideration previous recommendations delivered to the user. As such, the generated sequential recommendations will differ from a set of standalone recommendations, and this may be recognized by the user. In this scenario, explaining the dependencies between the recommended items or the reasons for recommending certain items may substantially improve user experience (Rana and Bridge 2020).

Sequential recommendations for groups One more open challenge relevant to sequential recommendations is group-based sequential recommendation. Group recommender system often build a virtual 'centroid' group member representing the group as a whole and deliver recommendations tailored to the preferences of this virtual group member. While this may turn out successful for homogeneous groups,

such recommendations may actually satisfy nobody in highly heterogeneous groups (Delic et al. 2016; Tran et al. 2018). On the contrary, the recommendation delivered by a sequential recommender could target a different member of the group every time and potentially improve the perceived quality of such recommendations. To this end, a framework for sequential group recommendations in the music domain was proposed by Hadash (2019). The key contribution of this framework was the satisfaction module that applied user and music track weighting to improve the fairness of the generated recommendations.

Improved offline evaluation Further improvements to our offline evaluation methodology are also possible. Today, research predominantly focuses on accuracy measures, such as the Mean Reciprocal Rank or the hit rate, and very few works consider other quality metrics, like coverage, diversity or novelty. In addition, certain recommendation algorithms might introduce undesirable biases, e.g., to recommend popular and fresh items, leading to reduced quality and potentially low uptake of the recommendations. In this sense, there might also exist longitudinal effects of certain recommendation strategies that are not well understood yet (Zhang et al. 2020). Generally, the community also seems to sometimes face methodological issues, e.g., comparisons with non-optimized baselines and limited reproducibility (Ferrari Dacrema et al. 2019; Ludewig et al. 2019), which may hamper the progress of research in session-based and sequential recommendations.

3 Papers in this issue

Using autoencoders for session-based job recommendations In this paper, Lacic et al. propose a session-based recommender system for job recommendation (Lacic et al. 2020). They are primarily concerned with a setting in which users anonymously interact with the system in online sessions, a scenario for which session-based techniques are well suited. The proposed approach uses several different autoencoder architectures, including denoising and variational autoencoders, to encode user sessions. The embeddings, in turn, are used in a k-nearest-neighbor recommendation framework to recommend jobs. The autoencoders are trained using two sources of data: interaction data from sessions, and content features of job postings, for which interactions were observed. The authors use three job recommendation data sets (two publicly available data sets and a proprietary data set from an Austrian student job portal) to train and test their models. They then report the results of several experiments measuring the performance of their approach using a variety of evaluation metrics. In particular, experimental results show that variational autoencoders that were trained on interaction and content data leads to the best results in terms of novelty and serendipity and good results in terms of accuracy compared to baselines.

Applying reranking strategies to route recommendation using sequence-aware evaluation Sánchez and Bellogín examine the problem of generating a meaningful and personalized sequence of venues (a route) for users in Location-Based Social Networks (LSBN) (Sánchez and Bellogín 2020). They develop a general framework for generating venue sequences from check-in data. The paper further investigates

a variety of reranking methods (across several standard recommendation models) to build sequences of items from recommendations. The paper also proposes a class of sequence-aware evaluation metrics (based on the Longest Common Subsequence algorithm) that take into account the ordering of visited venues in user profiles as part of the evaluation process. The experimental evaluation is performed using several LBSN datasets. The results generally show that those recommendation approaches that use temporal or geographical information perform better in terms of relevance. The experiments also explore various trade-offs among relevance, sequentiality, and distance, which are the key aspects in evaluating the quality of venue and route recommendations.

Session-based news recommendations using random walks on time-evolving heterogeneous networks In this paper, Symeonidis et al. develop a session-aware news recommender system which captures users' short-term intents by performing random walks on a heterogeneous information network (Symeonidis et al. 2020). The underlying time-evolving heterogeneous graph structure consists of users, news items and categories, sessions and other meta-data. The paper explores different sliding time windows and compares various weighting strategies as part of the session-aware recommendation generation process. The random walk with restart strategy is used to model the relevance of news items, while the sliding time windows are used to favor more recent articles in user profiles. The authors use three news data sets in their experimental evaluation. They empirically show that the proposed approach performs favorably when compared to several baselines. The best accuracy is achieved when considering very small session windows, for example, consisting of only the last item selected by the user. However, using longer session windows and the use of the heterogeneous information network may lead to more diverse recommendations.

References

- Berkovsky, S., Taib, R., Conway, D.: How to recommend? User trust factors in movie recommender systems. In: Proceedings of the 22nd International Conference on Intelligent User Interfaces (IUI'17), pp. 287–300 (2017)
- CIKM: The CIKM Cup 2016 (2016). <https://competitions.codalab.org/competitions/11161>. Accessed March 2020
- Delic, A., Neidhardt, J., Nguyen, T.N., Ricci, F., Rook, L., Werthner, H., Zanker, M.: Observing group decision making processes. In: Proceedings of the 10th ACM Conference on Recommender Systems (RecSys'16), pp. 147–150 (2016)
- Ekstrand, M.D., Harper, F.M., Willemsen, M.C., Konstan, J.A.: User perception of differences in recommender algorithms. In: Proceedings of the 2014 ACM Conference on Recommender Systems (RecSys'14), pp. 161–168 (2014)
- Ferrari Dacrema, M., Cremonesi, P., Jannach, D.: Are we really making much progress? a worrying analysis of recent neural recommendation approaches. In: Proceedings of the 2019 ACM Conference on Recommender Systems (RecSys) (2019)
- Hadash, S.: Evaluating a Framework for Sequential Group Music Recommendations. Master's thesis, Eindhoven University of Technology (2019)

- Hariri, N., Mobasher, B., Burke, R.: Context-aware music recommendation based on latent topic sequential patterns. In: Proceedings of the Sixth ACM Conference on Recommender Systems (RecSys'12), pp. 131 (2012)
- Hidasi, B., Quadrana, M., Karatzoglou, A., Tikk, D.: Parallel recurrent neural network architectures for feature-rich session-based recommendations. In: Proceedings of the 10th ACM Conference on Recommender Systems (RecSys'16), pp. 241–248 (2016)
- Jannach, D., Lerche, L., Jugovac, M.: Adaptation and evaluation of recommendations for short-term shopping goals. In: Proceedings of the 9th ACM Conference on Recommender Systems (RecSys'15), pp. 211–218 (2015)
- Jannach, D., Ludewig, M., Lerche, L.: Session-based item recommendation in e-commerce: On short-term intents, reminders, trends, and discounts. *User Model. User Adapt. Interact.* **27**(3–5), 351–392 (2017)
- Kamekhosh, I., Jannach, D.: User perception of next-track music recommendations. In: Proceedings of the 2017 Conference on User Modeling Adaptation and Personalization (UMAP'17), pp. 113–121 (2017)
- Knijnenburg, B.P., Willemsen, M.C., Gantner, Z., Soncu, H., Newell, C.: Explaining the user experience of recommender systems. *User Model. User Adapt. Interact.* **22**, 441–504 (2012)
- Lacic, E., Reiter-Haas, M., Kowald, D., Dareddy, M.R., Cho, J., Lex, E.: Encoding user sessions for online job recommendations. In: *User Modeling and User-Adapted Interaction* (2020)
- Li, J., Ren, P., Chen, Z., Ren, Z., Lian, T., Ma, J.: Neural attentive session-based recommendation. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM'17), pp. 1419–1428 (2017)
- Ludewig, M., Jannach, D.: User-centric evaluation of session-based recommendations for an automated radio station. In: Proceedings of the 2019 ACM Conference on Recommender Systems (RecSys'19), pp. 516–520 (2019)
- Ludewig, M., Mauro, N., Latifi, S., Jannach, D.: Performance comparison of neural and non-neural approaches to session-based recommendation. In: Proceedings of the 13th ACM Conference on Recommender Systems (RecSys'19), pp. 462–466 (2019)
- Mobasher, B., Dai, H., Luo, T., Nakagawa, M.: Using sequential and non-sequential patterns in predictive web usage mining tasks. In: Proceedings of IEEE International Conference on Data Mining (ICDM'02), pp. 669–672 (2019)
- Phuong, T.M., Thanh, T.C., Bach, N.X.: Neural session-aware recommendation. *IEEE Access* **7**, 86884–86896 (2019)
- Pu, P., Chen, L., Hu, R.: A user-centric evaluation framework for recommender systems. In: Proceedings of the 5th ACM Conference on Recommender Systems (RecSys'11), pp. 157–164 (2011)
- Quadrana, M., Karatzoglou, A., Hidasi, B., Cremonesi, P.: Personalizing session-based recommendations with hierarchical recurrent neural networks. In: Proceedings of the 2017 ACM Conference on Recommender Systems (RecSys'17) (2017)
- Quadrana, M., Cremonesi, P., Jannach, D.: Sequence-aware recommender systems. *ACM Comput. Surv.* **54**, 1–36 (2018)
- Ragno, R., Burges, C.J.C., Herley, C.: Inferring similarity between music objects with application to playlist generation. In: Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR'05), pp. 73–80 (2005)
- Rana, A., Bridge, D.: Navigation-by-preference: a new conversational recommender with preference-based feedback. In: Proceedings of the 25th International Conference on Intelligent User Interfaces (IUI'20), pp. 155–165 (2020)
- Resndle, S., Freudenthaler, C., Schmidt-Thieme, L.: Factorizing personalized Markov chains for next-basket recommendation. In: Proceedings of the 2010 World Wide Web Conference (WWW'10), pp. 811–820 (2010)
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: Grouplens: An open architecture for collaborative filtering of netnews. In: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW'94), pp. 175–186 (1994)
- Ricci, F., Del Missier, F.: Supporting travel decision making through personalized recommendation. In: Karat, C., Blom, J.O., Karat, J. (eds.) *Designing Personalized User Experiences in eCommerce*, Human-Computer Interaction Series, vol. 5, pp. 231–251. Kluwer, Dordrecht (2004)
- Rook, L., Sabic, A., Zanker, M.: Engagement in proactive recommendations. *J. Intell. Inf. Syst.* **54**(1), 79–100 (2020)

- Sánchez, P., Bellogín, A.: Applying reranking strategies to route recommendation using sequence-aware evaluation. In: *User Modeling and User-Adapted Interaction* (2020)
- Shani, G., Heckerman, D., Brafman, R.I.: An MDP-based recommender system. *J. Mach. Learn. Res.* **6**, 1265–1295 (2005)
- Symeonidis, P., Kirjackaja, L., Zanker, M.: Session-aware news recommendations using random walks on time-evolving heterogeneous information networks. In: *User Modeling and User-Adapted Interaction* (2020)
- Tavakol, M., Brefeld, U.: Factored MDPs for detecting topics of user sessions. In: *Proceedings of the 8th ACM Conference on Recommender Systems (RecSys'14)*, pp. 33–40 (2014)
- Tran, T.N.T., Atas, M., Felfernig, A., Samer, R., Stettinger, M.: Investigating serial position effects in sequential group decision making. In: *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP'18)*, pp. 239–243 (2018)
- Zhang, J., Adomavicius, G., Gupta, A., Ketter, W.: Consumption and performance: understanding longitudinal dynamics of recommender systems via an agent-based simulation framework. *Inf. Syst. Res.* **31**(1), 76–101 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dietmar Jannach is a professor of computer science at the University of Klagenfurt, Austria. His research is centered on practical applications of Artificial Intelligence, with a focus on recommender systems. He is also the leading author of the first textbook on recommender systems published with Cambridge University Press.

Bamshad Mobasher is a professor of Computer Science and the director of the Center for Web Intelligence at DePaul University. His research areas include Web mining, Web personalization, and recommender systems. He serves as an associate editor for the *ACM Transactions on the Web*, the *ACM Transactions on Internet Technology* and the *ACM Transactions on Intelligent Interactive Systems* and is on the editorial boards of several other prominent computing journals, including *User Modeling and User-Adapted Interaction*, and the *Journal of Web Semantics*.

Shlomo Berkovsky is the leader of the Precision Health research stream at Macquarie University. His work focuses on the use of machine learning methods to develop patient models and personalized predictions of diagnosis and care. He also studies how physiological responses can predict medical conditions, and how clinicians and patients interact with health technologies.