

Improving Session Based Recommendation by Diversity Awareness

Ramazan Esmeli¹, Mohamed Bader-El-Den¹, and Hassana Abdullahi¹

¹University of Portsmouth, Portsmouth PO1 2UP, UK,
{ramazan.esmeli,mohamed.bader,hassana.abdullahi}@port.ac.uk

Abstract. Recommender systems help users to discover and filter new and interesting products based on their preferences. Session-Based Recommender systems are powerful tools for anonymous e-commerce visitors to understand their behaviours and recommend useful products. Diversity in the recommendations is an important parameter due to increasing the opportunity of recommending new and less similar items that users interacted. Effect of diversity has been investigated in many works for the collaborative filtering-based Recommender systems. However, for session-based Recommender systems, exploring the effect of diversity is still an open area. In this paper, we propose an approach to calculate the diversity level of the items in the session logs and analyse the effect of diversity level on the session-based recommendation. In order to test the impact of diversity awareness, we propose a sequential Item-KNN recommendation model. The final recommendation list is created as a contribution of the interacted items in the session that depends on the diversity level between last interacted item of the session. We conduct several experiments to validate our diversity aware model on a real-world dataset. The results show that diversity awareness in the sessions helps to improve the performance of Recommender system in terms of recall and precision evaluation metrics. Also, the proposed method can be applied to other sequential Recommender system methods, including deep-learning based Recommender systems.

Keywords: Session Based Recommender Systems, Diversity, Context awareness

1 Introduction

Nowadays, due to the increase of available online data, users have many options in digital platforms, and they need to spend the effort to find the most relevant information for them. Information overloading, especially in the e-commerce domain, attracts researchers' and businesses' attention to develop approaches to filter most relevant products to their customers. Recommender Systems(RS) have been used more than a decade on digital platforms in different domains including movie, music and e-commerce[1] to help users to filter most relevant and preferred items for them.

RS are mainly divided into three categories based on the method used. These methods are Collaborative Filtering(CF) [18], Content-Based recommendation(CB)[4] and Hybrid RS[1]. In CF, user histories are used to train RS models, while in CB RS user and item attributes are utilised to create the models. However, Hybrid RS are built as a combination of CF and CB RS. There are drawbacks and advantages of these different recommendation methods. For example, for CF methods, user past behaviours are needed, for CB RS, user and item attributes need to be known, and Hybrid RS can take longer time and be computationally expensive.

Recently, new methods are emerging where there are missing user-interaction history and user-item attributes. For example, Session-Based Recommendation Systems (SBRS) emerge. In SBRS recommendations need to be given based on the only on-going session since users are anonymous and there is no other user-interaction history except the current session [13, 27, 15]. Many different methods have been developed to adapt to RS algorithms for anonymous users and short user-item interaction history including Item-Item KNN [22], Recurrent Neural Network(RNN) [13], Markov Chain(MC)[29], and various type of Matrix Factorisation (MF) approaches [17, 11]. Recent studies show that different version of RNN and modified Item-Item KNN methods[24] give promising results on SBRS domain in terms of recommendation quality by adapting to user intentions in the sessions and personalised recommendations.

The main aim of the developed methods in the SBRS domain is to improve the personalisation and user satisfaction by giving precise recommendations. Diversity in the recommendation area is a factor which can affect user satisfaction and recommendation quality[10]. In the literature, there are many works on diversity to analyse its effect on RS' performance[10, 19, 5]. However, in the SBRS domain, this factor is still an open area that needs to be investigated.

In this work, we propose Context and Diversity Aware (CDA) framework that measures the diversity on the on-going user session and adjust the next product recommendations based on the diversity level in the session. We evaluate the proposed framework on a real-world session based dataset. Computational results show that considering the interacted item diversity in the session improves the recommendation accuracy.

The main contributions of this paper are:

1. develop an approach to calculate diversity level of the items in the e-commerce sessions.
2. diversity level and session context integration into the SBRS framework.
3. validation of the proposed framework on a real-world dataset.

Rest of the paper is organized as follows. Section 2 discusses existing models in brief, which will be helpful in understanding SBRS, diversity and context awareness. Section 3 explains the proposed framework that identify the process through diversity measurement and context awareness. Also, this section will include the explanation of various steps used in the framework. Section 4 presents the data analysis for the dataset and the experiment results of the proposed

models. Section 5 presents the conclusion, recommendation and future scope for this work.

2 Related Works

This section gives an overview of SBRS, diversity and context awareness in the users' session logs.

2.1 Session Based Recommendation Systems

Traditional RS methods, for example, CF, CB and Hybrid heavily depend on user information and user history[1]. However, customers visiting the e-commerce platforms mostly prefer to browse anonymously[16]. In this case, traditional RS approaches suffer from cold start problem [21]. Therefore, traditional RS models cannot create recommendations since these models need to learn latent factors of the users as a vector to represent them[21].

Item-Item RS methods generally have been used when users are anonymous[14, 16] since item correlations are trained by looking at their co-occurrence in sessions. Amazon is one of the well-known examples of using item-item RS for SBRS[22]. Recently, Deep learning has been applied to SBRS domain in several works[14, 16, 13]. Their results show promising results in comparison to item-item RS. Also, in [16, 24], authors modified the item-item KNN RS to have more personalised and sequential aware model. Their experiment results on using same dataset showed that modified KNN models had better accuracy performance than RNN models.

MC based and combination of MF and MC based approaches[24, 17, 11] are also used in the case of missing user information. However, these models are computationally expensive compared to item-item KNN approaches[24].

2.2 Diversity in RS

Diversity is defined as the opposite of similarity. Moreover, diversification is introduced as one of the option to deal with over-fitting in recommendation[19]. Many diversity aware algorithms are proposed, and several workshops are organised to measure the effect of diversification on RS quality[19]. [9] examined the effect of the level of diversity on RS accuracy. Their results showed that most of the RS algorithms recommend fewer diversity products to increase RS accuracy.

Furthermore, [28] investigated novelty and diversity in the RS domain in terms of their effect on RS accuracy. The author calculated the diversity by taking into account the product position in the recommendation list and its similarity with other recommended items. Also, [5] created a measurement formulation of diversity in CB-RS by measuring the similarity of the given recommendations. They had a different recommendation list which has various level of diversity, and they conducted a user study. They found that when the diversity increases for a certain level, the user satisfaction improves. However, above the

identified level, users complained about seeing irrelevant recommendations to their preferences. Also, in [20], authors analysed the temporal diversity in CF, using user-user KNN and Singular Value Decomposition(SVD) methods. In their method, they trained the models using records from different time frames then they re-ranked the recommendation list based on the diversity rate between the most popular items. Also, [6] analysed diversity in the top-N RS. They re-ranked the recommendations by selecting the most diverse items to each other. Finally, [10] for the long tail problem(less popular items appears in the tail of list) in the recommendation system, they proposed an optimisation method that increases the diversity of RS. Their results show that the accuracy of RS is not affected; however, less popular items appeared in the top of the list.

Previous works [6, 20, 10] focused diversity in CF based Recommender systems. They mostly focused to diversity analysing on recommended items. Also, they re-ranked final recommendation list. However, in our work, we design a method that takes into diversity consideration before creating last recommendation list. Therefore, we do not apply re-ranking in final recommendation list.

2.3 Context Awareness in Session-Based Recommender Systems

User sessions are generally interpreted as a context in SBRS domain [27]. Users' interest shifting and browsed items in the session, the time difference between the clicked item, discounts on the item can be considered as the factors that build the session context[27, 16]. RNN based models can take into consideration the clicked items in the session and can give the next recommendation based on the clicked items. Also, RNN models can handle the popularity of the viewed items in the session as context[26]. Moreover, [15] adopted short-term user's intention to filter recommendations from various RS algorithms, including Bayesian Personalised Ranking(BPR) and Session-KNN. Also, paper[23] adopted CF recommendation for news broadcasting websites by using users' last intention and they re-ranked final recommendation list in terms of similarity measured by short term content features and recommended item features.

In our work, we use the users' short term intention(last viewed item in the session) to find the diversity level with previously seen products. So, in our work, we change the context contribution to final recommendation list based on diversity levels between last interacted item and the other item in the session. However, we do not apply re-ranking to recommendation list since created recommendation list will have already diversity awareness and ranked.

3 Proposed Approach

The proposed framework consists of 4 phases(Fig. 1). In this section, the phases are described.

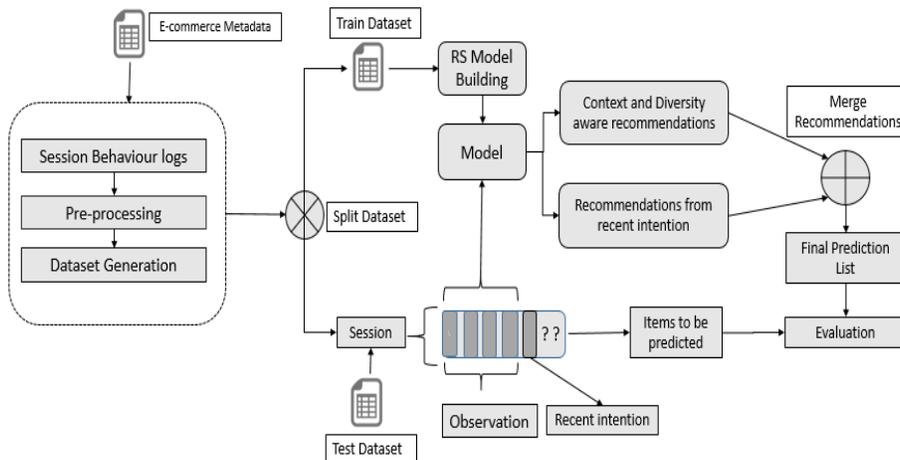


Fig. 1. Proposed Context and Diversity Aware SBRS Framework

3.1 Phase 1

In this phase, we prepare the session logs to fit RS algorithms. First, we select the related attributes from datasets, which are timestamp that shows when the session started, session id and item id, which are stored as string character in the dataset. We apply label encoding to make inputs suitable for RS algorithms. Also, in this phase, we clean the sessions which have less than 2 item interactions.

3.2 Phase 2

In this phase, we split the dataset as train and test. We use the last 2 days of the sessions as test dataset, and the others for train dataset. Also, in this phase, we train the Session-Item-Item KNN model[13].

3.3 Phase 3

One of the key contributions of this work is highlighted in this phase. In this phase, we build a diversity-aware SBRS, in which we use the diversity level of items between last interacted item to create a recommendation list(Fig. 2).

Diversity level calculation Diversity means in our context, the dissimilarity level between two products in a session. We follow a novel method to find the similarity of two items in the context of the given session to calculate their dissimilarity. In our method, first, we ask the trained recommendation model for n recommendation for each interacted product in the session and we create a recommendation list R_{i_p} . Later, we get n recommendation for the last interacted item in the session and create a recommendation list R_{i_l} . Finally we measure

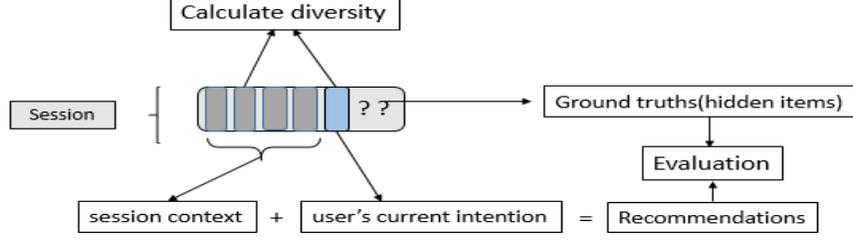


Fig. 2. Diversity calculation and its contribution to last recommendation list

the common recommended items between R_{i_p} and R_{i_l} . We keep common recommended items list in R_C (Eq. (1)).

$$R_C = R_{i_p} \cap R_{i_l} \quad (1)$$

So, similarity between these two items is calculated as in Eq.(2)

$$Sim_{(i_p, i_l)} = \frac{|R_C|}{n} \quad (2)$$

We have seen that when we include higher similar items with last interacted item in the session, the final recommendation list $FinalList$ in Eq. (4) will be almost the same with recommendation list R_{i_l} gathered from last interacted item. Therefore, overall, the contribution of the previously seen items will be less. However, when we give more contribution weight to recommendation list created from items which have more diversity level with the current intention, they can be ranked in the top- n recommendation list as their contribution scores are increased. So we calculate a diversity-aware contribution weight for the previously seen items in the session as in Eq. (3)

$$W_p = 1 - Sim_{(i_p, i_l)} \quad (3)$$

In Eq.(4), the final recommendation list is calculated. T is the number of interacted items in the session s . I is the whole items in the train dataset. $Score_p$ and $Score_l$ are confident scores of top- n neighbours of item p and item l , which are given from the recommendation model.

$$Final\ List = top_n\left(\sum_p^{T-1} W_p * Score_p(i_p, I)\right) + top_n(Score_l(i_l, I)) \quad (4)$$

In the Eq. (4), we choose n parameter as 100 to select the most 100 (n) similar *neighbours*. In this phase, n is used to calculate common recommended items between last interacted item and previously seen items. However as seen in Eq. (4), we do not give a contribution weight to the last interacted item.

3.4 Phase 4

In this phase, we evaluate the performance of our approach using $Recall@n$ and $Precision@n$ metrics[12, 24]. n means in the evaluation context is selecting top- n high scored items in the ranked final recommendation list(Eq. (4)). In this work, we choose $n \in \{10, 20, 30\}$ since users are mostly interested in the recommendations in the top of the list.

4 Dataset Analysing

In this section, we give statistical details of the dataset we use in this paper. The dataset collected for a two weeks period from an e-commerce platform in the UK(Fresh relevance¹). Before pre-processing the dataset, most of the sessions have a few item interactions as seen in Fig. 3. We eliminate the sessions which

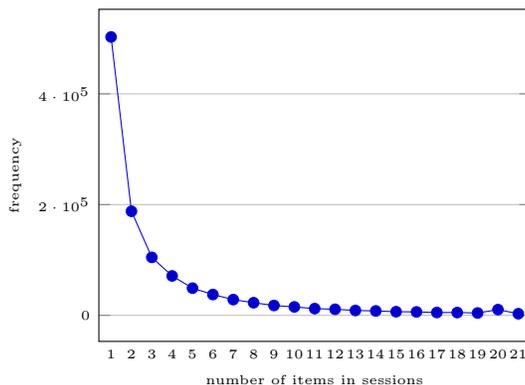


Fig. 3. Item frequency in the sessions

have less than 2 items interaction to deal with cold start problem. Also, having more interactions in the session can help to create better similarity correlations between items.

Before processing and after processing the dataset statistics is shown in Table 1

Table 1. Dataset Statistics

dataset	Sessions	Items	Interactions
fresh relevance before	961653	42363	2844362
fresh relevance after	466818	37537	2349527

¹ <https://www.freshrelevance.com/>

5 Experimental Evaluation

We conduct four experiments to see the effect of integration of diversity to SBRS.

Experiment 1 In the first experiment, we create recommendations from last interacted item only. So, the given recommendation from this model does not take into consideration session context. We call this model as Item-Item KNN (Fig 4).

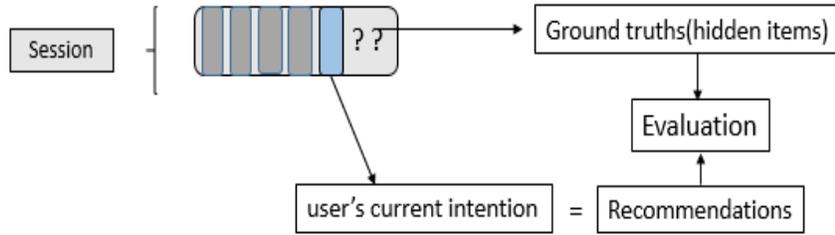


Fig. 4. Only user's current intention and its contribution to last recommendation list

Experiment 2 In this experiment, we consider the contribution of the session context. We call this model as C-Item KNN (Fig. 5).

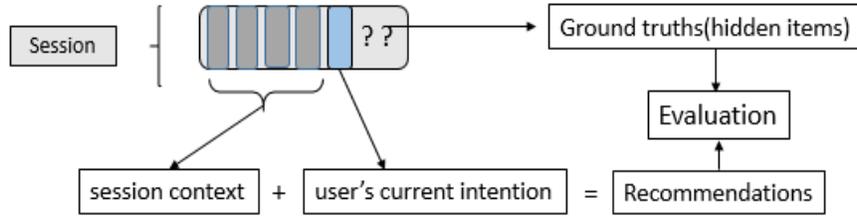


Fig. 5. Only Session context awareness and its contribution to last recommendation list

Experiment 3 The aim of this experiment is to see effect of diversity awareness in the context on SBRS performance. This model has interaction and diversity awareness. We call this model CD-Item KNN (Fig. 2).

Experiment 4 We aim in this experiment to see the effect of decreasing diversity on recommender system quality. In other words, more similar items to last interacted item will have more contribution weight. So, we updated Eq. (3) as in Eq. (5).

$$W_p = Sim(i_p, i_l) \quad (5)$$

In this final experiment, the last recommendation list will be created as a consequence of similar item interactions. We call this model SI-Item KNN.

Evaluation Method To evaluate the experiments we hide a part of the interacted items from the session. The final recommendation list created in each experiment is ranked. From the ranked list, we choose the top- n ($n \in \{10, 20, 30\}$) best ranked recommendations since users mostly interested in recommendations which are in the top of the list. The model tries to predict hidden interacted items from the session which we call ground truths. We use the $Recall@n$ (Eq. (6)) and $Precision@n$ (Eq. (7)) metrics to measure how the models performing on predicting missing interactions in the session.

$$Recall@n = \frac{|groudtruth \cap recommendations@n|}{|groudtruth|} \quad (6)$$

$$Precision@n = \frac{|groudtruth \cap recommendations@n|}{|recommendations@n|} \quad (7)$$

We tested the models with 1000 sessions. To choose test sessions, we iterated selection process ten times over the test dataset and in each iteration, 100 sessions are randomly selected from dataset. Overlapping of new selected 100 sessions with previously selected sessions are prevented to minimize the risk that the obtained results are specific to the single test split as shown in Figure 6. Finally, the performance measurements of all splits are averaged.

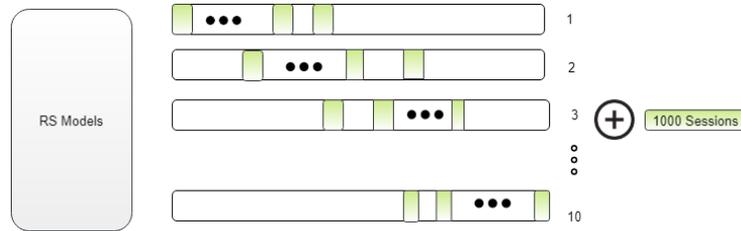


Fig. 6. Selecting test sessions to validate CDA framework

5.1 Results and Discussion

The experiment results (Table 2) show that adding diversity awareness to the contribution of the interacted items in the session context leads to good performance in terms of Recall and Precision in comparison to when diversity and

context awareness are not included in the models. Moreover, the results confirm that context awareness in the RS domain, particularly in SBRS is an essential factor to have better-performing models. Increasing the diversity level of the items to build the final recommendation list helps by considering the different intentions of the user in the session. Moreover, remembering items which are diverse to users current intention and similar to session context help to have improve recommendations.

On the other hand, experiment result confirms that when the weight of contribution of similar interacted items with last interacted item is increased, the performance of SBRS becomes even worse than created recommendations from only last interacted item. Base on literature[19, 5], CF-based RS show opposite performance results than this experiment. The reason for our result could be as a result of the specification of SBRS since in the sessions, users have a specific target. Therefore, browsed items will have a level of similarity. However, remembering the less similar products that are already seen in the session can match user behaviour in the session. Thus, more diversity level to users' last intention increases the RS performance.

Table 2. Performance Comparision of Models on Rec:Recall, Prec:Precision @ $n \in \{10, 20, 30\}$

Model	Rec@10	Prec@10	Rec@20	Prec@20	Rec@30	Prec@30
Item-Item KNN	0.2386	0.0239	0.3143	0.0157	0.3571	0.0119
C-Item KNN	0.3171	0.0317	0.4186	0.0209	0.4700	0.0157
DC-Item KNN	0.3514	0.0351	0.4371	0.0219	0.4800	0.0160
SI-Item KNN	0.2171	0.0217	0.2986	0.0149	0.3386	0.0113

6 Conclusion and Future Works

In this paper, we applied diversity adjustment to session context to create SBRS recommendation list. For diversity calculation, we developed a novel equation and proposed an algorithm to integrate it into the SBRS model. We applied the algorithm on session based dataset. In order to test the proposed algorithm, we created an efficient diversity and context-aware(CDA) SBRS framework. In this framework, interacted item contribution is increased when diversity increases between previously interacted item and last interacted item in a session to build up final recommendation list. Also, we give the highest contribution to last interacted item to define the final recommendation list since last interacted item will show a higher level of user intention than previously seen items[15, 27]. Also, we used Item-Item KNN and modified Item-Item KNN models in the framework for our experiments. Our results showed that integrating dynamic item diversity in SBRS can improve recommendation performance in terms of precision and recall metrics.

For future works, the proposed diversity algorithm can be applied to other SBRS models, including deep learning, BPR, MC and Association Rule-based models. Also, the diversity level can be restricted to a certain level. For instance, items which have above a certain level of item diversity with users' current intention can be disregarded while building the final recommendation list. Lastly, we will test our algorithm on other session-based datasets to see the performance of our approach.

Another future direction is that calculated diversity can be used in machine learning classifications, including text classification [25] which can help to predict feature specifications of products from the user search query that the user might click or browse in the session. However, in these type of classification approaches, class imbalance is the main challenge [3, 2] and to process data in real-time and in parallel, tools such as Apache Spark [7, 8], need to be considered, in which recommendation process need to be done within a few seconds.

References

1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge & Data Engineering* (6), 734–749 (2005)
2. Bader-El-Den, M.: Self-adaptive heterogeneous random forest. In: 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA). pp. 640–646. IEEE (2014)
3. Bader-El-Den, M., Teitei, E., Perry, T.: Biased random forest for dealing with the class imbalance problem. *IEEE transactions on neural networks and learning systems* (2018)
4. Besbes, O., Gur, Y., Zeevi, A.: Optimization in online content recommendation services: Beyond click-through rates. *Manufacturing & Service Operations Management* 18(1), 15–33 (2015)
5. Castagnos, S., Brun, A., Boyer, A.: When diversity is needed... but not expected! In: International Conference on Advances in Information Mining and Management. pp. 44–50. IARIA XPS Press (2013)
6. Di Noia, T., Ostuni, V.C., Rosati, J., Tomeo, P., Di Sciascio, E.: An analysis of users' propensity toward diversity in recommendations. In: Proceedings of the 8th ACM Conference on Recommender systems. pp. 285–288. ACM (2014)
7. Diedhiou, C., Carpenter, B., Esmeli, R.: Comparison of platforms for recommender algorithm on large datasets. In: 7th Imperial College Computing Student Workshop. pp. 4–1. Schloss Dagstuhl–Leibniz Center for Informatics (2019)
8. Diedhiou, C., Carpenter, B., Shafi, A., Sarkar, S., Esmeli, R., Gadsdon, R.: Performance comparison of a parallel recommender algorithm across three hadoop-based frameworks. In: 2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD). pp. 380–387. IEEE (2018)
9. Fleder, D.M., Hosanagar, K.: Recommender systems and their impact on sales diversity. In: Proceedings of the 8th ACM conference on Electronic commerce. pp. 192–199. ACM (2007)
10. Hamedani, E.M., Kaedi, M.: Recommending the long tail items through personalized diversification. *Knowledge-Based Systems* 164, 348–357 (2019)

11. He, R., McAuley, J.: Fusing similarity models with markov chains for sparse sequential recommendation. In: 2016 IEEE 16th International Conference on Data Mining (ICDM). pp. 191–200. IEEE (2016)
12. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)* 22(1), 5–53 (2004)
13. Hidasi, B., Karatzoglou, A., Baltrunas, L., Tikk, D.: Session-based recommendations with recurrent neural networks. arXiv preprint arXiv:1511.06939 (2015)
14. 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. pp. 241–248. ACM (2016)
15. 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. pp. 211–218. ACM (2015)
16. Jannach, D., Ludewig, M.: When recurrent neural networks meet the neighborhood for session-based recommendation. In: Proceedings of the Eleventh ACM Conference on Recommender Systems. pp. 306–310. ACM (2017)
17. Kabbur, S., Ning, X., Karypis, G.: Fism: factored item similarity models for top-n recommender systems. In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 659–667. ACM (2013)
18. Karabadjji, N.E.I., Beldjoudi, S., Seridi, H., Aridhi, S., Dhifi, W.: Improving memory-based user collaborative filtering with evolutionary multi-objective optimization. *Expert Systems with Applications* 98, 153–165 (2018)
19. Kunaver, M., Požrl, T.: Diversity in recommender systems—a survey. *Knowledge-Based Systems* 123, 154–162 (2017)
20. Lathia, N., Hailes, S., Capra, L., Amatriain, X.: Temporal diversity in recommender systems. In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. pp. 210–217. ACM (2010)
21. Lika, B., Kolomvatsos, K., Hadjiefthymiades, S.: Facing the cold start problem in recommender systems. *Expert Systems with Applications* 41(4), 2065–2073 (2014)
22. Linden, G., Smith, B., York, J.: Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing* (1), 76–80 (2003)
23. Liu, J., Dolan, P., Pedersen, E.R.: Personalized news recommendation based on click behavior. In: Proceedings of the 15th international conference on Intelligent user interfaces. pp. 31–40. ACM (2010)
24. Ludewig, M., Jannach, D.: Evaluation of session-based recommendation algorithms. *User Modeling and User-Adapted Interaction* 28(4-5), 331–390 (2018)
25. Mohasseb, A., Bader-El-Den, M., Cocea, M.: Question categorization and classification using grammar based approach. *Information Processing & Management* 54(6), 1228–1243 (2018)
26. Moreira, G.d.S.P., Jannach, D., da Cunha, A.M.: Contextual hybrid session-based news recommendation with recurrent neural networks. arXiv preprint arXiv:1904.10367 (2019)
27. Quadrana, M., Cremonesi, P., Jannach, D.: Sequence-aware recommender systems. *ACM Computing Surveys (CSUR)* 51(4), 66 (2018)
28. Vargas, S.: New approaches to diversity and novelty in recommender systems. In: Fourth BCS-IRSG symposium on future directions in information access (FDIA 2011), Koblenz. vol. 31 (2011)
29. Wang, W., Yin, H., Sadiq, S., Chen, L., Xie, M., Zhou, X.: Spore: A sequential personalized spatial item recommender system. In: 2016 IEEE 32nd International Conference on Data Engineering (ICDE). pp. 954–965. IEEE (2016)