COMP9945 Research Proposal

Structure-aware Interactive Recommendation

September 22, 2019

1 Introduction

Recommender systems is one of the most successful applications of machine learning technology in real world. It designed to give users certain suggestions within large range of objects in a personalized way. It has been widely used in many application domains, such as e-commerce, web navigation recommadation, media streaming and become a inevitable form of service of people daily online user expericence.

Regarding the developing process of recommender system, the traditional recommender syster can be considered as a matrix completion problem that is trying to predict the missing value of a given user-item rating matrix. However, one limitation could be the cold start problem which is a typical issue for matrix factorization. Moreover, this method is only suited to longer-term cases that is collecting data during a long time to capture longer-term users preference profiles.

It has been proved that the short-term user histories can deeply affect the next decision making of users. For example, in online shopping, people who recently changed his shopping interest (say from digital devices to fashion cloth) can have a different set of perefernce during a short time. Those histories are not limited to preference and purchase records. It can be views, clicks, queries etc which are more of interactive actions. In addition, for some certain problem e.g. session-based recommendation from Hidasi et al. [5] which longer-term log is not available. For example, there might consistently be some new-registered or anonymous user which we do not have any record for. Therefore, short-term user histories can be central to a success recommender system and need to be considered into applications. Such scenario can be characterized as a sequence-aware recommendation problem.

In practice, most of the algorithmic approach of sequence-aware recommendation procedure as a static process make recommendations following a fixed greedy strategy. These methods might fail to give user suggestions in a interatively way and fit the users dynamic perefernce. In some cases such as music domain, it is important to keep listening the feed back from user then give a dynamic reaction. That is reason, in this research, we will leverage reinforcement learning to better model this problem. Therefore the whole recommender sysytem are capable of interacting with people and considering both long-term and shor-term behaviors.

2 Review of the literature

Regarding the sequence-aware recommendation, there are a variety of research going on focusing on different problem tasks. Previous works in this field are relatively scattered. According to Quadrana, Cremonesi, and Jannach [11], all attention on this problem can be classified as four categorizations regarding sequence-aware recommendation tasks:

1. Context Adaptation
2. Trend Detection
3. Repeated Redcommendation
4. Consideration of Order Constarints and Sequential Patterns

In terms of Context Adaptation, we believe the recommendable item can depend on the users current sitiation e.g. location, weather, people around as well, such as Cheng et al. [2]. While for the Trend Detection task, researchers were dedicated to detect the change of interest at individual or the variation of popularity of the items, Jannach and Ludewig [6]. Thirdly, Repeated Redcommendation aims to detect repeated user behavior patterns, Lu, Lin, and Ciou [9]and/or as reminder, Lerche, Jannach, and Ludewig [8]which can be meaningful for user expericence. Finally, there are some works resolving the problem of ordering constrains e.g. Its not a good to recommand a latest series of movie for a person who did not watch preious one, Parameswaran, Venetis, and Garcia-Molina [10].

In addition, we can review the current existed work regarding methodology. As the input of a sequence-aware recommendation should be an ordered list of past actions, when some researchers implement it, the action does not need to be related to a item. For example, some behaviors such as search, navigation can also be considered, according to Jannach and Ludewig [7].

In the feature engineering part, a common item embedding method is being used. Greenstein-Messica, Rokach, and Friedman [4] used item embedding which encode every product with one-hot encoding then combine items together within a session. This method is similar to the world embedding in NLP which can let us take the order of the actions into account.

Algorithms is another important aspect of the implementation. We have found most of the efforts based on the Sequence Learning methods which is a common choice for pattern mining problem. In particular, there are some significant works based on RNN, Markov chains, RL. Zhang et al. [15] use RNNs for click prediction for online advertisements.In their thesis, they train the RNN to predict the next click of the users given their last click and the previous state of the network using a classification loss measure (cross-entropy). Based on the RNN structure, there are few exploration on using Gated Recurrent Units (GRUs) for modeling user activity in a session-based scenario Chung et al. [3]. While one of the constrains of using RNN is that the computation can not be parallelized, for the large size input it can be time-consuming. Instead, Zhang et al. [14] used self-attention to tackle this problem. Chen et al. [1] went another way, they consider user behaviors as neural turing machines and introduce a memory mechanism to the recommander systems which achieve better performance than the former models.

However, as discuss in the section 1, these methods lack interaction with users. In the case of users having changable interest, reinforcenment learning (RL) could be a effective solution. Zheng et al. [17] use a Deep Q-Learning based recommendation framework, which can model future reward explicitly. While when the number of items increased, a Deep Q-Learning or POMDP (Ricci, Rokach, and Shapira [12]) method would become inflexible. To address this problem, some recommender systems are build on a A3C RL which have improvement on large item scale. In addition, with the benifit of the concept of reward in RL, the researchers such as Shani, Brafman, and Heckerman [13] have made some bonus. They not only predict the best next item for user but also the income(total reward) for the saler.

For a complete recommender system, especially a interactive system, online test is a key part for the final performance. Zhao et al. [16] build a online user-agent interacting environment simulator, which is suitable for offline parameters pre-training and evaluation before applying a recommender system online in the real world.

Overall, sequence-aware interactive recommendation scenarios are highly relevent in practice and many significant works were proposed in the past. However, disadvantages had been found in both two areas. Therefore, it is crucial to come up with a novel recommender sysyterm frame structure which leads to the aim of this research discussed in the next section.

3 Research aims

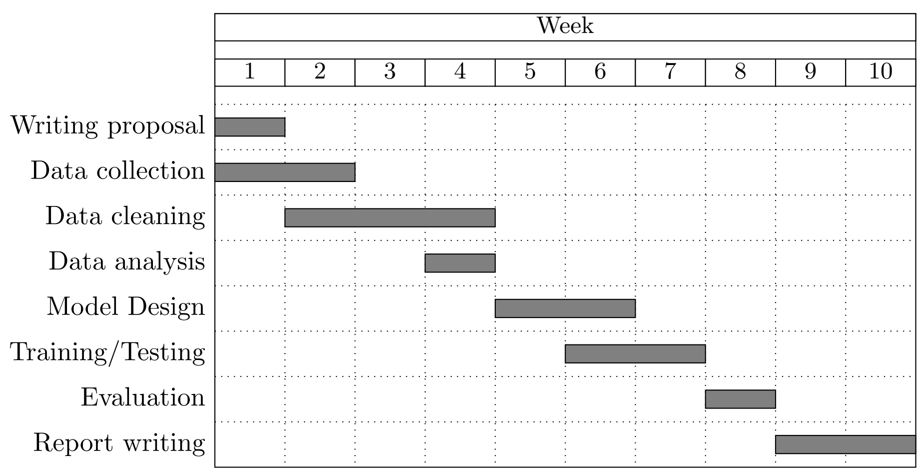
Limitations of current recommender system. (1) Traditional recommender system does not take short-term interaction behavior into account which leads to lack of latest perference of users. (2) Most of the methods do not interact with user which causes the lack of dynamic interest detection of users. (3) Majority implementation of RL recommends/outputs only one item with highest Q-values, which lost the diversity of a person's interest, for example, in the News application, a new user might expect a large range of topics to choose, instead of a similar article based on the past click. Meanwhile, it does not give users a structured-sequence recommendation which could benifit user experience in some application such as Online-course recommendation.

*Aims and objectives.*To tackle the limitations above. This research project proposal a novel recommender system with the combination of sequence-aware recommendation and reinforcenment learning. In the implementation, it would be using public datesets and a ordered sequential action from user log as input, item embedding and a RL structure based on A3C. Output a list of high-Q-value items as recommendation to predict a series structured-sequence recommendation for users. Finally evaluate the system performance by common metrics such as MAP or NDCG. Expecting a better porformance compared with traditional methods especially in Media, Online-shopping, News domains.

Challengs. (1) One potential challeng would be data cleaning, as we are getting sequential actions of user, we have to integrate the data so that it can be used as input. (2) Secondlly, Its hard to determine the five elements of a Markov Decision Process (MDP) problem while implementing RL. (3) Some Dataset might potentially not fit this framework.

Potential reactions. (1) More techniques would be using when it comes to data cleaning. It could use numpy/pandas restructured the raw data to achieve the correct form for training data. (2) As we are dealling with a situation where a recommender agent interacts with users by sequentially choosing recommendation items over a sequence of time steps, so as to maximize its cumulative reward, once we define the reward other elements could be figured out naturally. One possible defination is that if user click or order other items amongst the output list, we could get the model a positive reward as feedback. (3) Dataset selection would be mainly focused on Music(spotify), E-commerce, News application domains.

4 Timeline



References

[1] Xu Chen et al. "Sequential Recommendation with User Memory Networks". In: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining.* WSDM '18. Marina Del Rey, CA,

USA: ACM, 2018, pp. 108–116. ISBN: 978-1-4503-5581-0. DOI: 10.1145/3159652.3159668. URL: http://doi.acm.org/10.1145/3159652.3159668.

[2] Chen Cheng et al. "Where You Like to Go Next: Successive Point-of-interest Recommendation". In: *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*. IJCAI '13. Beijing, China: AAAI Press, 2013, pp. 2605–2611. ISBN: 978-1-57735-633-2. URL: http://dl.acm.org/citation.cfm?id=2540128.2540504.

[3] Junyoung Chung et al. "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling". In: *ArXiv* abs/1412.3555 (2014).

[4] Asnat Greenstein-Messica, Lior Rokach, and Michael Friedman. "Session-Based Recommendations Using Item Embedding". In: *Proceedings of the 22Nd International Conference on Intelligent User Interfaces*. IUI '17. Limassol, Cyprus: ACM, 2017, pp. 629–633. ISBN: 978-1-4503-4348-0. DOI: 10.1145/3025171.3025197. URL: http://doi.acm.org/10.1145/3025171.3025197.

[5] Balázs Hidasi et al. "Session-based Recommendations with Recurrent Neural Networks". In: *CoRR* abs/1511.06939 (2015).

[6] Dietmar Jannach and Malte Ludewig. "Determining Characteristics of Successful Recommendations from Log Data: A Case Study". In: *Proceedings of the Symposium on Applied Computing*. SAC '17. Marrakech, Morocco: ACM, 2017, pp. 1643–1648. ISBN: 978-1-4503-4486-9. DOI: 10.1145/3019612.3019757. URL: http://doi.acm.org/10.1145/3019612.3019757.

[7] Dietmar Jannach and Malte Ludewig. "Investigating Personalized Search in E-Commerce". In: *FLAIRS Conference*. 2017.

[8] Lukas Lerche, Dietmar Jannach, and Malte Ludewig. "On the Value of Reminders Within E-Commerce Recommendations". In: *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. UMAP '16. Halifax, Nova Scotia, Canada: ACM, 2016, pp. 27–35. ISBN: 978-1-4503-4368-8. DOI: 10.1145/2930238.2930244. URL: http://doi.acm.org/10.1145/2930238.2930244.

[9] E. H. Lu, Y. Lin, and J. Ciou. "Mining mobile application sequential patterns for usage prediction". In: *2014 IEEE International Conference on Granular Computing (GrC)*. Oct. 2014, pp. 185–190. DOI: 10.1109/GRC.2014.6982832.

[10] Aditya Parameswaran, Petros Venetis, and Hector Garcia-Molina. "Recommendation Systems with Complex Constraints: A Course Recommendation Perspective". In: *ACM Trans. Inf. Syst.* 29.4 (Dec. 2011), 20:1–20:33. ISSN: 1046-8188. DOI: 10.1145/2037661.2037665. URL: http://doi.acm.org/10.1145/2037661.2037665.

[11] Massimo Quadrana, Paolo Cremonesi, and Dietmar Jannach. "Sequence-Aware Recommender Systems". In: *CoRR* abs/1802.08452 (2018). arXiv: 1802.08452. URL: http://arxiv.org/abs/1802.08452.

[12] Francesco Ricci, Lior Rokach, and Bracha Shapira. "Introduction to Recommender Systems Handbook". In: *Recommender Systems Handbook*. 2011.

[13] Guy Shani, Ronen I. Brafman, and David Heckerman. "An MDP-based Recommender System". In: *CoRR* abs/1301.0600 (2013). arXiv: 1301.0600. URL: http://arxiv.org/abs/1301.0600.

[14] Shuai Zhang et al. "Next Item Recommendation with Self-Attention". In: *CoRR* abs/1808.06414 (2018). arXiv: 1808.06414. URL: http://arxiv.org/abs/1808.06414.

[15] Yuyu Zhang et al. "Sequential Click Prediction for Sponsored Search with Recurrent Neural Networks". In: *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*. AAAI'14. Québec City, Québec, Canada: AAAI Press, 2014, pp. 1369–1375. URL: http://dl.acm.org/citation.cfm?id=2893873.2894086.

[16] Xiangyu Zhao et al. "Deep Reinforcement Learning for List-wise Recommendations". In: *ArXiv* abs/1801.00209 (2017).

[17] Guanjie Zheng et al. "DRN: A Deep Reinforcement Learning Framework for News Recommendation". In: Apr. 2018, pp. 167–176. DOI: 10.1145/3178876.3185994.