Conversational Recommendation: Formulation, Methods, and Evaluation

Lei, W.; He, X.; de Rijke, M.; Chua, T.-S.

DOI
10.1145/3397271.3401419

Publication date
2020

Document Version
Author accepted manuscript

Published in
SIGIR '20

Citation for published version (APA):
Lei, W., He, X., de Rijke, M., & Chua, T.-S. (2020). Conversational Recommendation: Formulation, Methods, and Evaluation. In *SIGIR '20: proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval : July 25-30, 2020, virtual event, China* (pp. 2425–2428). Association for Computing Machinery. https://doi.org/10.1145/3397271.3401419

Conversational Recommendation: Formulation, Methods, and Evaluation

Wenqiang Lei
National University of Singapore
wenqianglei@gmail.com

Xiangnan He
University of Science and Technology of China
hexn@ustc.edu.cn

Maarten de Rijke
University of Amsterdam & Ahold Delhaize
derijke@uva.nl

Tat-Seng Chua
National University of Singapore
dcscts@nus.edu.sg

ABSTRACT

Recommender systems have demonstrated great success in information seeking. However, traditional recommender systems work in a static way, estimating user preferences on items from past interaction history. This prevents recommender systems from capturing dynamic and fine-grained preferences of users. Conversational recommender systems bring a revolution to existing recommender systems. They are able to communicate with users through natural languages during which they can explicitly ask whether a user likes an attribute or not. With the preferred attributes, a recommender system can conduct more accurate and personalized recommendations.

Therefore, while they are still a relatively new topic, conversational recommender systems attract great research attention. We identify four emerging directions: (1) exploration and exploitation trade-off in the cold-start recommendation setting; (2) attribute-centric conversational recommendation; (3) strategy-focused conversational recommendation; and (4) dialogue understanding and response generation. This tutorial covers these four directions, providing a review of existing approaches and progress on the topic.

By presenting the emerging and promising topic of conversational recommender systems, we aim to provide take-aways to practitioners to build their own systems. We also want to stimulate more ideas and discussions with audiences on core problems of this topic such as task formalization, dataset collection, algorithm development, and evaluation, with the ambition of facilitating the development of conversational recommender systems.

CCS CONCEPTS

• **Information systems → Users and interactive retrieval**; **Recommender systems**; *Personalization*; • **Human-centered computing** → *Interactive systems and tools.*

KEYWORDS

Conversational recommendation, Dialog system, Interactive recommendation

SIGIR '20, July 25–30, 2020, Virtual Event, China

ACM ISBN 978-1-4503-8016-4/20/07.
https://doi.org/10.1145/3397271.3401419

ACM Reference Format:
Wenqiang Lei, Xiangnan He, Maarten de Rijke, and Tat-Seng Chua. 2020. Conversational Recommendation: Formulation, Methods, and Evaluation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3397271.3401419

1 INTRODUCTION

Recommender systems have become the prevalent choice for information seeking. Companies such as Amazon and Alibaba, in e-commere, Facebook and Wechat, in social networking, and Instagram and Pinterest, in content sharing, all aim to link items (products, posts, etc.) to users in an effective manner with the goal of matching user preferences. An effective recommendation with both accuracy and timeliness can help users find desired information, greatly save their time, and bring significant value to business. Therefore, recommendation techniques have attracted continuously growing academic and industrial attention.

Traditional recommender systems, which we call **static recommender systems** in this tutorial, primarily predict a user's preference towards an item by analyzing their past behavior, e.g., click history, visit log, ratings on items, etc. Sophisticated methods have been proposed and proven to be effective. For example, (neural) factorization machines [6, 21] and deep interest networks [31] make use of historical user-item interactions to estimate users' preferences based on the collaborative filtering hypothesis, which assume that similar users may have similar interests. Those supervised models have achieved great success by being trained on offline data and deployed in many real-world scenarios. Recently developed models [25, 26] further use the expressiveness and explainability of graphs to model complex relations among users, items and attributes, with the aim of better modeling such offline data.

However, the static recommendation paradigm has a fundamental intrinsic limitation: it is not able to capture a user's current preferences, which may have shifted away from their historical ones. In addition, it is hard to find accurate reasons as to why a user interacts with certain items as there is no channel for static recommendation system to access such information. **Conversational recommender systems** (CRSs) [3–5, 14, 18, 19, 22–24, 28, 30] promise to bring a solution to such problems. With CRSs, recommendation becomes an interactive process through which users converse with the recommender using natural language. In this manner, a CRS should be able to obtain or infer the dynamic preferences of users from their utterances, which might consist of their direct or indirect descriptions of their needs or of answers to questions posed by the recommender system in a mixed-initiative setup. Such information can help a CRS to adjust its recommendation strategy effectively and, we hope, dramatically, enhance its performance.

2 OVERVIEW

Motivated by the potential of conversational recommender systems, recent years have witnessed extensive research efforts being devoted to exploring this topic. However, such efforts spread various task formulations which are under different hypotheses and application scenarios. For the purposes of this tutorial, we categorize them in four directions, and plan to discuss their formulations, methods and evaluations. The four directions are as follows:

2.1 Exploitation-Exploration Trade-offs for Cold Users

This direction is derived from bandit online recommendation methods [13, 15]. Work on conversational recommendations takes advantage of such bandit approaches to balance the *exploration and exploitation* trade-offs for cold-start users in conversational recommendation scenarios [5, 29]. Christakopoulou et al. [5] make the first step to adapt a typical bandit-based method like UCB and Thompson Sampling [2] to the conversational recommendation scenario. Sepliarskaia et al. [23] learn to ask preference questions to help profile a cold user. Zhang et al. [29] develop UCB methods to consider user feedback on attributes, where an attribute is a description of the properties of items.[1] Following this line, Li et al. [16] build a ConTS model based on Thompson Sampling for cold start users.

2.2 Question Driven Approaches

A key feature of conversational recommendation system is its ability to chat with users. With this capability, a CRS can directly ask questions to users to get more information about their preferences. While this property is addressed as "asking clarification question" in generally conversational information seeking [1], researchers studying conversational recommendation pays specific attention to the attribute. They believe it is effective to proactively consult a user whether he or she likes an attribute. Therefore, the estimation and utilization of a user's preferences towards attributes become a key research issue of asking attribute-centric question. Zhang et al. [30] use a multi-memory network to estimate user preferred attributes to consult the user from their profile and product description in an e-commerce scenario. Yu et al. [28] aim to understand user comments on a batch of recommended items guide the model for next round of recommendation. Christakopoulou et al. [4] propose a two-stage paradigm, demonstrating potential attributes (referred to as *topics* in the original paper) for users to select before making a recommendation.

[1]For example, being an item, a restaurant is given with its location, favor, average score, etc. as attributes. Note that there are no uniform terminologies in the topic of conversational recommendation systems as it is quite new. An *attribute* is referred to as a *key-term* in Zhang et al. [29] while being referred to as a *facet* in [24].

2.3 Multi-turn Conversational Recommendation Strategy

A CRS might have multi-turn conversations with users. During each turn, a CRS can ask questions about the user's preferences (usually asking the user's preferences on different attributes) or make recommendations. Intuitively, the clarification questions about the user preference can lead to better recommendation quality. However, if a CRS asks too many questions, the lengthy conversation will lead to users being bored. Therefore, a CRS needs to strategically ask questions and make recommendations to achieve engaging and successful recommendations with fewer turns. Work along this line includes [10, 12, 24].

2.4 Dialogue Understanding and Generation

As core problems of dialogue systems, natural language understanding and generation are crucial in CRSs. This direction focuses on how to understand a user's preferences and intentions from their utterances and generate fluent responses so as to deliver natural and effective dialogue actions (e.g., asking question and making recommendations) [8, 9, 11, 20]. Multiple datasets and simulation environments have been released to help push the statee-of-the-art in this area For instance, Li et al. [14] release a dataset comprising of more than 10,000 dialogues on movie recommendation and Chen et al. [3] incorporate a knowledge graph to bridge dialogue understanding and generation with the recommendation component. Li et al. [17] use simulations to train adversarial methods to generate better dialogue responses. And Liao et al. [18] approach conversational recommendation on travel domains, emphasizing the importance of topic managing.

The rapid development of recommender systems and dialogue agents in recent years provides us with an opportunity to help define how these two technologies can be combined. Now is the time to activate the community around this topic. Given that significant progress is being made, on many different aspects of conversational recommendation systems, we believe it is timely to review and bring together what has done so far. We hope to stimulate more ideas as well as discussions to promote the development of this emerging and promising research topic further, building on the rich tradition that the information retrieval community can bring to the table for this topic, in terms of algorithm development, understanding user preferences, human-computer interaction, and evaluation.

3 OBJECTIVES

By offering this tutorial, we want to benefit both industry and academia:

(1) We aim to provide industrial participants with a broad picture of the current development of conversational recommenders, with key take ways for building realistic conversational recommendation for their own scenario.
(2) We aim to provide academic participants with comprehensive literature review and insightful discussions. By summarizing existing assumptions and explorations on the topic of conversational recommendation, together with a review of recent progresses on recommender systems and dialogue

systems, we hope to engage in deeper discussion with the audience, sparking ideas for core problems for this topic. For example, how to fuse or integrate different research directions, how to better leverage on existing efforts of recommendation and dialogue systems, how to better formalize tasks and conduct evaluation for conversational recommendation.

In particular, we want to convey the following in this tutorial:

- **Share advances in static conversational recommendation techniques.**
  We review traditional static recommender systems by introducing representative models like early factorization methods [21], neural collaborative filtering [7] as well as the graph-based recommendation [25–27]. We review their key assumption, strengths and shortcomings, and analyze the need to develop conversational recommendation.
- **Summarize recent developments in dialogue systems.**
  We cover the development of task-oriented dialogue system and non-task-oriented dialogues (a.k.a. chitchat system and chatbot). Specifically, we will review four typical components for dialogue systems: the natural language understander, the belief tracker, the policy maker and the response generation. We discuss how those techniques might be potentially helpful to conversational recommendation.
- **Identify, describe, and discuss different research directions for conversational recommendations.**
  This is the main part of this tutorial. We plan to introduce the four categories of work sketched above, with a particular emphasis on their application scenarios, methods, datasets and evaluations. Based on that, we aim to distil core research questions on this topic. Based on that, we want to engage in discussions with the participants about the future research.

4 FORMAT AND DETAILED SCHEDULE

The tutorial is organized into four parts, with the aim to discuss how recommender systems and dialogue systems are brought together to compose conversational recommender systems. In Part I, we give an introduction, viewing the topic of conversational recommendation from the development of information retrieval technology perspective. In Part II and III, we give preliminaries of recommender systems and dialogue systems. This does not only provide necessary background regarding both research topics, but also provides inspiration for why and how these two topics can help each other. In Part IV, the core of the tutorial, we provide a detailed introduction to existing efforts on building conversational recommender systems, using the four categories described in Section 1.

The following summarizes the schedule of the tutorial, with timing:

I: Introduction (10 min)
- 1.1 Conversational recommendation from information retrieval
- 1.2 Organization of the tutorial

II: Preliminaries of recommender systems (20 min)
- 2.1 Collaborative filtering
- 2.2 Deep learning approaches
- 2.3 Graph-based approaches

III: Preliminaries of dialogue systems (20 min)
- 3.1 Task-oriented dialogue systems
- 3.2 Chit-chat dialogue systems
- 3.3 Template-based and Seq2seq dialogue systems

IV: Conversational recommendation systems (120 min)
- 4.1 Exploitation-exploration balance for cold-users
- 4.2 Attribute-centric conversational recommendation
- 4.3 Strategy-focused conversational recommendation
- 4.4 Dialogue understanding and generation

V: Future directions (10 min)

5 SUPPORTING MATERIALS

We will share the following materials with the participants of the tutorial:

(1) **Slides** All slides will be made publicly available.
(2) **Annotated bibliography** An annotated compilation of references will list all works discussed in the tutorial and should provide a good basis for further study.
(3) **Code** An annotated list of pointers to open source code and datasets for the work discussed in our tutorial will be shared with attendees.
(4) **Survey** The authors are writing a survey on conversational recommendation; a complete draft will be shared with attendees.

The materials can be found at https://core-tutorial.github.io.

ACKNOWLEDGMENTS

Maarten de Rijke is partially supported by the Innovation Center for AI (ICAI). This research is also supported by the National Research Foundation, Singapore under its International Research Centres in Singapore Funding Initiative as well as National Natural Science Foundation of China (61972372, U19A2079). All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

REFERENCES

[1] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W. Bruce Croft. 2019. Asking Clarifying Questions in Open-Domain Information-Seeking Conversations. In *SIGIR*. SIGIR, 475–484.
[2] Olivier Chapelle and Lihong Li. 2011. An Empirical Evaluation of Thompson Sampling. In *NeurIPS*. 2249–2257.
[3] Qibin Chen, Junyang Lin, Yichang Zhang, Ming Ding, Yukuo Cen, Hongxia Yang, and Jie Tang. 2019. Towards Knowledge-Based Recommender Dialog System. In *EMNLP-IJCNLP*. Association for Computational Linguistics, 1803–1813.
[4] Konstantina Christakopoulou, Alex Beutel, Rui Li, Sagar Jain, and Ed H Chi. 2018. Q&R: A Two-Stage Approach toward Interactive Recommendation. In *SIGKDD*. 139–148.
[5] Konstantina Christakopoulou, Filip Radlinski, and Katja Hofmann. 2016. Towards Conversational Recommender Systems. In *SIGKDD*. 815–824.
[6] Xiangnan He and Tat-Seng Chua. 2017. Neural factorization machines for sparse predictive analytics. In *SIGIR*. 355–364.
[7] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *WWW*. 173–182.
[8] Shaojie Jiang, Pengjie Ren, Christof Monz, and Maarten de Rijke. 2019. Improving Neural Response Diversity with Frequency-Aware Cross-Entropy Loss. In *The Web Conference 2019*. ACM, 2879–2885.
[9] Xisen Jin, Wenqiang Lei, Zhaochun Ren, Hongshen Chen, Shangsong Liang, Yihong Zhao, and Dawei Yin. 2018. Explicit State Tracking with Semi-Supervisionfor Neural Dialogue Generation. In *CIKM*. ACM, 1403–1412.
[10] Wenqiang Lei, Xiangnan He, Yisong Miao, Qingyun Wu, Richang Hong, Min-Yen Kan, and Tat-Seng Chua. 2020. Estimation–Action–Reflection: Towards Deep Interaction Between Conversational and Recommender Systems. In *WSDM*. ACM, 304–312.
[11] Wenqiang Lei, Xisen Jin, Min-Yen Kan, Zhaochun Ren, Xiangnan He, and Dawei Yin. 2018. Sequicity: Simplifying Task-oriented Dialogue Systems with Single Sequence-to-Sequence Architectures. In *ACL*. 1437–1447.

[12] Wenqiang Lei, Gangyi Zhang, Xiangnan He, Yisong Miao, Xiang Wang, Liang Chen, and Tat-Seng Chua. 2020. Interactive Path Reasoning on Graph for Conversational Recommendation. In *Proceedings of the 2020 ACM SIGKDD Conference on Knowledge Discovery and Data Mining*.

[13] Lihong Li, Wei Chu, John Langford, and Robert E Schapire. 2010. A Contextual-bandit Approach to Personalized News Article Recommendation. In *WWW*. ACM, 661–670.

[14] Raymond Li, Samira Ebrahimi Kahou, Hannes Schulz, Vincent Michalski, Laurent Charlin, and Chris Pal. 2018. Towards Deep Conversational Recommendations. In *NeurIPS*. 9748–9758.

[15] Shuai Li, Alexandros Karatzoglou, and Claudio Gentile. 2016. Collaborative Filtering Bandits. In *SIGIR*. 539–548.

[16] Shijun Li, Wenqiang Lei, Qingyun Wu, Xiangnan He, Peng Jiang, and Tat-Seng Chua. 2020. Seamlessly Unifying Attributes and Items: Conversational Recommendation for Cold-Start Users. *arXiv preprint arXiv:2005.12979* (2020).

[17] Ziming Li, Julia Kiseleva, and Maarten de Rijke. 2019. Dialogue Generation: From Imitation Learning to Inverse Reinforcement Learning. In *AAAI 2019: 33rd AAAI Conference on Artificial Intelligence*. AAAI.

[18] Lizi Liao, Yunshan Ma, Xiangnan He, Richang Hong, and Tat-Seng Chua. 2018. Knowledge-aware Multimodal Dialogue Systems. In *ACM MM*. 801–809.

[19] Bilih Priyogi. 2019. Preference Elicitation Strategy for Conversational Recommender System. In *WSDM*. ACM, 824–825.

[20] Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2020. Thinking Globally, Acting Locally: Distantly Supervised Global-to-Local Knowledge Selection for Background Based Conversation. In *Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)*. AAAI.

[21] Steffen Rendle. 2010. Factorization Machines. In *ICDM*. 995–1000.

[22] Nicola Sardella, Claudio Biancalana, Alessandro Micarelli, and Giuseppe Sansonetti. 2019. An Approach to Conversational Recommendation of Restaurants. In *ICHCI*. Springer, 123–130.

[23] Anna Sepliarskaia, Julia Kiseleva, Filip Radlinski, and Maarten de Rijke. 2018. Preference Elicitation as an Optimization problem. In *RecSys 2018: The ACM Conference on Recommender Systems*. ACM, 172–180.

[24] Yueming Sun and Yi Zhang. 2018. Conversational Recommender System. In *SIGIR*. 235–244.

[25] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural Graph Collaborative Filtering. In *SIGIR*. 165–174.

[26] Xiang Wang, Dingxian Wang, Canran Xu, Xiangnan He, Yixin Cao, and Tat-Seng Chua. 2019. Explainable Reasoning over Knowledge Graphs for Recommendation. In *AAAI*, Vol. 33. 5329–5336.

[27] Yikun Xian, Zuohui Fu, S Muthukrishnan, Gerard De Melo, and Yongfeng Zhang. 2019. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation. In *SIGIR*. 285–294.

[28] Tong Yu, Yilin Shen, and Hongxia Jin. 2019. An Visual Dialog Augmented Interactive Recommender System. In *SIGKDD*. ACM, 157–165.

[29] Xiaoying Zhang, Hong Xie, Hang Li, and John Lui. 2020. Conversational Contextual Bandit: Algorithm and Application. In *WWW*.

[30] Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W Bruce Croft. 2018. Towards Conversational Search and Recommendation: System Ask, User Respond. In *CIKM*. 177–186.

[31] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep Interest Network for Click-through Rate Prediction. In *KDD*. 1059–1068.