Lessons Learnt from Building Friend Recommendation Systems

Jun Yu
jyu3@snapchat.com
Snap Inc.
Santa Monica, CA, USA

ABSTRACT

Friend recommendation systems in online social networks such as Snapchat help users find friends and build meaningful connections, leading to heightened user engagement and retention. While friend recommendation systems follow the classical recommendation system paradigm that consists of retrieval and ranking, they pose distinctive challenges different from item recommendation systems (e.g. Youtube videos, Amazon products, Netflix movies), and require special considerations in building one. In this paper, we elucidate the unique challenges encountered and share invaluable insights from developing the friend recommendation system for hundreds of millions of users on Snapchat.

CCS CONCEPTS

• **Information systems** → **Learning to rank**;

KEYWORDS

Friend Recommendation, Social Networks, User Retrieval, Graph Neural Network

ACM Reference Format:
Jun Yu. 2024. Lessons Learnt from Building Friend Recommendation Systems. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining (WSDM '24), March 4–8, 2024, Merida, Mexico.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3616855.3635750

1 FRIEND RECOMMENDATION SYSTEMS

In online social and professional networks [1, 9], a user's friends or connections are critical for one's engagement and retention. Research at Linkedin [15] showed that "members with at least 13 connections from companies other than their current employer are 22.9% faster in transitioning to their next job than those who do not". Friend recommendation is often formulated as an industrial recommendation problem and follows a typical large-scale recommendation system architecture [3] which consists of two stages: *retrieval* and *ranking*. Figure 1 demonstrates the friend recommendation funnel where candidates are first retrieved and then ranked before getting surfaced to the end users. This two-stage recommendation system advantageously sidesteps quadratic complexity from considering all user-user pairs, while also enabling flexible trade-offs in recall, precision and infrastructure cost.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
WSDM '24, March 4–8, 2024, Merida, Mexico
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0371-3/24/03.
https://doi.org/10.1145/3616855.3635750

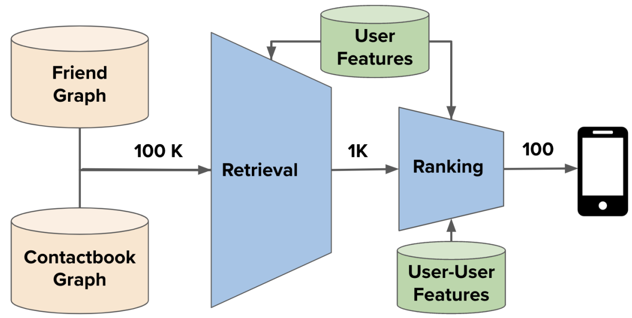

Figure 1: Friend Recommendation Architecture.

2 CHALLENGES AND LEARNINGS

Despite sharing a similar paradigm as item recommendations, friend recommendations pose unique challenges throughout the system.

2.1 Retrieval

In the *retrieval* phase, we extract friend candidates using one's friend graph and contactbook graph. The number of candidates is often in the hundreds of thousands. These candidates are then ranked and filtered based on either heuristics or lightweight machine learning models, and around top ten-thousand candidates are funneled to the next phase of ranking. The goal of the retrieval phase is to include as many high potential friends as possible and recall is used to measure its effectiveness.

Friend of Friend (FoF). In item recommendations, user-user and item-item collaborative filtering [5] is the most commonly used approach in retrieval. Graph traversal approaches (e.g. FoF [2]) have been used extensively in retrieving FoF candidates and are adopted widely in online social and professional networks such as LinkedIn [10] and Facebook [6]. This approach extracts candidates using breadth-first traversal and is often built on an inverted index for fast querying [11].

Embedding Based Retrieval. Embedding based retrieval is an effective retrieval model [7, 13] in both item and friend recommendations. It requires pre-computed embeddings and support fast queries based on embedding similarities. In item recommendations, two tower model [14] is a prominent approach to generate item and user embeddings in a transductive manner. In friend recommendations, we found embeddings from inductive learning (e.g. Graph Neural Networks [12]) more effective empirically. Embeddings learnt from these two learning paradigms complement each other, indicating different information being captured in each.

Cold Start. Cold start problems in item recommendations can be either on new users or new items. For new users, we recommend popular items based on their demographics. For new items, we promote them using their similarities to existing items or metadata such as brand, genre, etc. However, friend recommendations are highly personalized and the aforementioned approaches fail to work effectively. Instead, we rely on users' contactbook and their connections to identify the first set of their potential friends.

2.2 Ranking

In the *ranking* phase, we rank retrieved candidates using deep neural network models and send around top 100 friend recommendations to the end users. The ML model is trained on historical friend recommendation outcomes and uses both user-level features and user-to-user interactions as signals to rank candidates. The goal of the ranking phase is to maximize precision.

Ranking Objective. In item recommendations, the goal is to optimize user engagement such as click, purchase, share, etc, which are explicitly modeled in the ranking objective. In friend recommendations, we need to model the entire friending funnel from the initial friend request to quality of friendship. These events are conditioned on their up-streams, e.g. reciprocation of a friend request is required since friendship is bidirectional. To optimize quality friendship, we model each event along the funnel as well as the interaction depth using multi-task learning.

Position Bias. In video recommendations, there is a strong position bias with which users tend to click on the top recommended videos. Researchers have studied position bias and developed debiasing approaches [16]. In friend recommendations, we didn't observe strong position bias for typical users because the "cost" of adding a wrong friend is much higher than watching a less interesting video.

Privacy Requirement. Unlike item recommendations, different privacy requirements in social networks are appropriate and can be imposed to friend recommendations which require separate considerations. For example, friend lists of users on Snapchat are private by design. In compliance with this privacy requirement, we use differential privacy [4] to blend suggestions from different retrieval sources in the final ranked list such that one cannot easily infer another user's friend list.

2.3 Evaluation

Measuring the true rollout impact of a feature in friend recommendations is challenging due to network effect [8]. In an AB test, the same user can be included in the control group as a searcher and in the treatment group as a candidate. Another challenge is the AB results sometimes can not be linearly extrapolated in the full rollout. For example, the retention improvement by promoting a small set of churned users in the treatment will not scale linearly because they compete for the same real estate when being rolled out to all churned users. To address the challenges, we developed an AB framework that supports both candidate-side AB and searcher-side AB to get a more accurate estimate of a feature's full rollout impact.

ACKNOWLEDGMENTS

We thank the members of the Friending team for their dedication and innovations, Alexandre Rezende, Jiabin Chen, and Jie Liu for their invaluable data insights, Ios Kotsogiannis and Yuchao Tao for their support in privacy matters, Nik Srivastava and David Boyle for their partnership in the product, and Lei Shi and Ganesh Venkataraman for their leadership support.

COMPANY PORTRAIT

Snap Inc. is a technology company. We contribute to human progress by empowering people to express themselves, live in the moment, learn about the world, and have fun together.

PRESENTER BIO

Jun Yu is a senior engineering leader at Snap, building products that help users create meaningful connections on Snapchat. Prior to joining Snap, he held senior positions in applied machine learning at Amazon and eBay, contributing to a wide spectrum of machine learning applications. These included areas such as paid internet marketing, promotion pricing and scheduling, notification campaign optimization, and risk management. He holds a Ph.D. in Computer Science from Oregon State University and has extensively published in top-tier machine learning conferences and journals.

REFERENCES

[1] Danah M Boyd and Nicole B Ellison. 2007. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication* 13, 1 (2007), 210–230.

[2] Jilin Chen, Werner Geyer, Casey Dugan, Michael Muller, and Ido Guy. 2009. Make New Friends, but Keep the Old: Recommending People on Social Networking Sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA). 201–210.

[3] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*. New York, NY, USA.

[4] Cynthia Dwork and Aaron Roth. 2014. The Algorithmic Foundations of Differential Privacy. *Found. Trends Theor. Comput. Sci.* 9, 3–4 (aug 2014), 211–407.

[5] Michael D. Ekstrand, John T. Riedl, and Joseph A. Konstan. 2011. Collaborative Filtering Recommender Systems. *Found. Trends Hum.-Comput. Interact.* (2011).

[6] Facebook. 2023. *Where do People You May Know suggestions come from on Facebook?* https://www.facebook.com/help/163810437015615

[7] Jui-Ting Huang, Ashish Sharma, Shuying Sun, Li Xia, David Zhang, Philip Pronin, Janani Padmanabhan, Giuseppe Ottaviano, and Linjun Yang. 2020. Embedding-Based Retrieval in Facebook Search. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2553–2561.

[8] Michael L Katz and Carl Shapiro. 1985. Network Externalities, Competition, and Compatibility. *American Economic Review* 75, 3 (June 1985), 424–440.

[9] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Predicting Positive and Negative Links in Online Social Networks. In *Proceedings of the 19th International Conference on World Wide Web* (Raleigh, North Carolina, USA). Association for Computing Machinery, New York, NY, USA, 641–650.

[10] Linkedin. 2023. *People You May Know Feature.* https://www.linkedin.com/help/linkedin/answer/a544682/people-you-may-know-feature?lang=en

[11] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval.* Cambridge University Press, Cambridge, UK. http://nlp.stanford.edu/IR-book/information-retrieval-book.html

[12] Aravind Sankar, Yozen Liu, Jun Yu, and Neil Shah. 2021. Graph neural networks for friend ranking in large-scale social platforms. In *Proceedings of the Web Conference 2021*. 2535–2546.

[13] Jiahui Shi, Vivek Chaurasiya, Yozen Liu, Shubham Vij, Yan Wu, Satya Kanduri, Neil Shah, Peicheng Yu, Nik Srivastava, Lei Shi, Ganesh Venkataraman, and Jun Yu. 2023. Embedding Based Retrieval in Friend Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Taipei, Taiwan). 3330–3334.

[14] Xinyang Yi, Ji Yang, Lichan Hong, Derek Zhiyuan Cheng, Lukasz Heldt, Aditee Ajit Kumthekar, Zhe Zhao, Li Wei, and Ed Chi (Eds.). 2019. *Sampling-Bias-Corrected Neural Modeling for Large Corpus Item Recommendations.*

[15] Guillaume Saint Jacques YinYin Yu and Paul Matsiras. 2021. *How to build an effective professional network on LinkedIn: Some data-driven insights.* Retrieved April 2, 2021 from https://engineering.linkedin.com/blog/2021/professional-network-checklist

[16] Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed Chi. 2019. Recommending What Video to Watch next: A Multitask Ranking System. In *Proceedings of the 13th ACM Conference on Recommender Systems*. 43–51.