# Recommendation in Context-Rich Environment: An Information Network Analysis Approach

Yizhou Sun
University of California, Los Angeles
yzsun@cs.ucla.edu

Hongzhi Yin
The University of Queensland, Australia
db.hongzhi@gmail.com

Xiang Ren
University of Illinois at Urbana-Champaign
xren7@illinois.edu

## ABSTRACT

Recommendation has received tremendous attention recently due to its wide and successful applications across different domains. Different from traditional setting of recommendation tasks, modern recommendation tasks are usually exposed in a context-rich environment. For example, in addition to a user-item rating matrix, users and items are connected to other objects via different relationships and they are usually associated with rich attributes, such as text and spatio-temporal information. It turns out that heterogeneous information network serves a natural data model to capture the rich context of these recommendation tasks. In this tutorial, we will systematically introduce the methodologies of using heterogeneous information network mining approach to solve recommendation tasks, and demonstrate the effectiveness of such methods using different applications, ranging from collaboration recommendation in scientific research network to job recommendation in professional social network, and to drug discovery in biomedical networks. The topics to be covered in the tutorial include: (1) overall introduction; (2) recommendation in heterogeneous information networks, which introduces the general methodology of how to model the recommendation problem as a heterogeneous information network mining problem; (3) recommendation in a text-rich setting, where the information network is further enriched by refined analysis of text information; (4) recommendation with spatio-temporal information, where entities and relationships in the network are associated with spatio-temporal attributes; and (5) research frontiers for context-rich recommendation.

## Keywords

Recommender Systems; Information Network Mining; contex-rich recommendation; Spatio-temporal recommendation

## 1. ORGANIZERS

**Yizhou Sun** (Ph.D., UIUC), is an assistant professor at University of California, Los Angeles (UCLA). Her principal research interest is in mining information and social networks, and more generally in data mining, machine learning, and network science. She has over 60 publications, over 10 of which are on recommendation algorithms. She received 2012 ACM SIGKDD Best Student Paper Award, 2013 ACM SIGKDD Doctoral Dissertation Award, 2013 Yahoo! ACE (Academic Career Enhancement) Award, and 2015 NSF CAREER Award. Dr. Sun has rich experiences in delivering tutorials in major conferences, including EDBT 2009, SIGMOD 2010, KDD 2010, ICDE 2012, VLDB 2012, ASONAM 2012, and ACL 2015.

**Xiang Ren** is a fifth-year Ph.D. candidate of Department of Computer Science at University of Illinois at Urbana-Champaign (UIUC). His research focuses on creating generic computational tools for better understanding massive text data. He has published over 25 papers in major conferences. He received 2016 Google PhD Fellowship in Structured Data and Database Management, 2016 KDD Rising Star by Microsoft Academic Search, C. W. Gear Outstanding Graduate Student Award in 2016, Yahoo!-DAIS Research Excellence Award in 2015, and C. L. and Jane W.-S. Liu Award in 2015. Mr. Ren has rich experiences in delivering tutorials in major conferences, including SIGKDD 2015, SIGMOD 2016 and WWW 2016.

**Hongzhi Yin** (Ph.D., Peking University), works as an ARC DECRA Fellow (Lecturer-Level) with The University of Queensland (UQ), Australia. His main research interests are in social media analysis and recommender system, especially spatio-temporal recommendation. He has published over 30 papers, with more than 10 papers and a scholar book on spatio-temporal recommendation. He received 2016 Australia Discovery Early Career Researcher Award and 2014 Distinguished Doctor Degree Thesis Award of Peking University. Dr. Yin has been the lecturer for the course "Information Retrieval and Web Search" in the University of Queensland and delivered many oral presentations and tutorials in major conferences.

## 2. INTRODUCTION

Recommendation has received tremendous attention recently due to its wide and successful applications across different domains. Different from traditional setting of recommendation tasks, modern recommendation tasks are usually exposed in a context-rich environment. Take a movie recommendation task as an example, in addition to a user-movie rating matrix, users and movies are usually connected to other objects via different relationships. For example, users are linked to other users via friendship links, and movies are


*WWW'17 Companion,* April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3051105

linked to actors/actresses, directors, and genres. The rich context can be naturally captured by heterogeneous information networks, which contain multiple types of objects and their interactions. In other words, the user-movie interaction matrix is embedded into a more complicated heterogeneous information network. By using information network mining approaches, the rich information can be fully utilized for the recommendation task, which turns out to be extremely helpful in overcoming the sparsity and cold start issue in traditional recommendation methodologies.

Moreover, objects and relationships in the network are usually associated with rich text and spatio-temporal information. For example, in a citation recommendation task for scientific research, text information in papers become very critical to determine whether a paper should be considered as a reference. In another example of restaurant recommendation for users in a location-based social network service, spatio-temporal information become critical in determining the quality of the recommendation. In these cases, information networks will be further enriched by these types of information. We will then introduce how to integrate the text and spatio-temporal information into the information network, and how to utilize them to provide a better recommendation.

In this tutorial, we will introduce the recent progress in recommendation algorithms that are related to information network mining approach in an organized way, which will provide researchers and practitioners in this field with a new angle.

### Duration and Sessions.

This is one-day tutorial. The topics to be covered in the tutorial include: (1) overall introduction; (2) recommendation in heterogeneous information networks, which introduces the general methodology of how to model the recommendation problem as a heterogeneous information network mining problem; (3) recommendation in a text-rich setting, where the information network is further enriched by refined analysis of text information; (4) recommendation with spatio-temporal information, where entities and relationships in the network are associated with spatio-temporal attributes; and (5) research frontiers for context-rich recommendation.

### Audience and Prerequisite.

The tutorial is suitable for academic and industrial researchers, graduate students, and practitioners who are interested in dealing with recommendation tasks with rich contexts including social, textual, and spatio-temporal attributes. The audience will benefit from the new perspective of solving recommendation problems in an information network mining approach and the broad applications covered in the tutorial. Only basic knowledge of data mining, information networks, and recommender systems are needed.

### Tutorial Material.

The tutorial materials will be provided to attendees via: `http://web.cs.ucla.edu/~yzsun/Tutorials/WWW2017/index.html`.

## 3. TUTORIAL OUTLINE

The outline of the proposed tutorial is summarized below.

**Part I: Preliminaries.** We first introduce the preliminaries of recommendation and information network mining.

**Part II: Recommendation in Heterogeneous Information Networks.** Heterogeneous information network that contains multiple types of objects and multiple types of relationships serves a natural data model in capturing the rich context for many recommendation tasks. In this part, we introduce three types of methodologies that can leverage the rich semantics in heterogeneous information networks for recommendation.

1. **Recommendation as link prediction in heterogeneous information networks.** Recommendation tasks can be viewed as link prediction tasks in heterogeneous information network, where meta-path-based proximity features [25] can be formed for each source-target pair in the network and the link prediction problem is converted to classification problem. We introduce two examples to illustrate the methodology, including collaboration recommendation in scientific research network [26] and drug recommendation for protein target in drug discovery research [4].
2. **Recommendation with social network as side information.** Recent studies [27, 16, 17, 18] have found that social relations that are usually captured in social networks are very helpful for the recommendation tasks. For example, friends may share similar interests and thus like similar movies. Social network information can be considered as an additional type of links that connect users, in addition to the standard user-item interactions.
3. **Entity recommendation with implicit feedback via meta-path propagation.** In many cases, only implicit feedback is received for the recommendation task. In [44], the user-item implicit feedback information is designed to be propagated via meta-paths that are derived from heterogeneous information networks. Each meta-path implies an intention why a user likes an item. This leads to multiple models corresponding to different latent spaces. The best weighted combination of these latent spaces can be learned using Bayesian ranking optimization. In [42, 41], a personalized model that learns a separate set of weights for each user group is further proposed.
4. **Entity similarity regularization for recommendation as feature selection.** When facing rich context information, one of the biggest challenges is to select the most relevant information for the recommendation task. In [43], a multi-view graph-regularized matrix factorization framework is proposed to incorporate and learn the weights of heterogeneous relationships.

**Part III: Recommendation in a Text-Rich Setting.** In the context-rich environment, there exists not only different kinds of relationships between users, items and objects of various types, but also a rich amount of text information (e.g., in the forms of short-text descriptions, documents, keywords, etc.) attached to users, items and other objects. It is of great interests to study: (1) how to overcome the cold-start problem with the help of text information, and (2) how to incorporate rich text signals to improve the recommendation in heterogeneous information networks. In this part of the tutorial, we give an overview of the problems and methodologies in content-based recommendation, introduce the approaches for recommendation in text-rich information networks, and discuss about the methods for constructing structured networks from text data and performing recommendation based on the text networks so constructed.

1. **Content-based recommendation: An Overview.** We will introduce the basics of item representation and user representation and different methods for learning user and item representations, by using literature search as an example. We then provide an overview of the state-of-the-art recommendation models [1, 15] including keyword-based models and ontology-based models. In particular, we will discuss how user feedback can be leveraged in the models.
2. **Recommendation in text-rich information networks.** This part of the tutorial focuses on the cases where we have structured relationships organized as information networks (e.g., user-item interactions, item-attribute associations, and user-user connections), as well as text information attached to different objects in the networks (e.g., item description, user profile). We first introduce methods which build a *global recommendation models* in text-rich information network [19, 2, 45, 40, 14]. We discuss how to leverage text information to group users and derive *personalized recommendation model* for each user group [20]. To facilitate different ranking hypotheses, we introduce *restricted meta-paths* for recommendation in text-rich networks [13].
3. **Recommendation in networks constructed from text.** As the majority of existing data generated in our society is unstructured, we present studies that extract structures from text data to facilitate recommendation. We introduce *embedding techniques for weighted term networks* constructed from the corpora [9] for content-based recommendation. We also discuss methods which can extract *typed entities from text corpora* and leverage them to enrich the information networks [22, 23, 24, 21].

**Part IV: Recommendation with Spatio-Temporal Information.** The recent advances in location-acquisition and wireless communication technologies such as GPS have greatly promoted the rapid prevalence of smart mobile devices, which enables users to easily add a location dimension to traditional social networks (e.g., Twitter, Facebook, and Weibo) via a smartphone, and also fosters the growth of Location-based Social Networking services (LBSNs) such as Foursquare, Yelp, GeoLife, Meetup, and Google Place, where users can easily check-in at points of interests (e.g., restaurants, stores, hotels) and share their life experiences in the physical world via mobile devices, resulting in rich user-generated spatiotemporal data. Actually, the user-generated spatiotemporal data in the mobile era captures the snapshots of users' everyday lives, and provides unprecedented potential for user interest modeling, mobility pattern mining and user behavior prediction. On the other hand, it is crucial to develop spatiotemporal recommendation services for users to explore the new places, attend new events and find their potentially preferred items from billions of candidate ones.

Spatiotemporal recommendation aims to find top-$k$ spatial items (i.e., items with geographical locations) that a target user $u$ potentially prefers in a given specific spatiotemporal context (i.e., $q = (u, t, l)$) [34, 38]. According to the recent studies [39, 33, 46], human spatiotemporal behaviors are complex and influenced by multiple factors: *Geographical Influence*, *Temporal Cyclic Effect*, *Temporal Sequential Influence* and *Social Influence*. While there are some previous studies [36, 11, 37, 5, 7, 6] that exploit one of the above factors to improve spatiotemporal recommendation, they lack an integrated analysis of their joint effect. To capture the joint effect of all the above factors in a unified way, we will present two types of spatiotemporal recommendation methods based on the heterogenous information networks.

1. **Spatiotemporal recommendation by probabilistic generative models**. To mimic user spatiotemporal behaviors in a process of decision making, we will present a joint probabilistic generative model [39, 34, 38] and a sparse additive generative model [30, 31, 10] which strategically integrate the multiple factors mentioned above in a unified way. Compared with traditional mixture models, these two models do not need to introduce additional latent variables which act as "switches" to control the weights of each factor.
2. **Spatiotemporal recommendation by network embedding**. A heterogenous information network embedding method will be presented to learn the representations of users, spatial items, time slots, geographical regions and content words in a shared latent space [33, 8, 47], based on which we investigate the joint power of the above mentioned factors to address the challenges of data sparsity, cold start, context awareness and dynamic user preferences in spatiotemporal recommendation.

**Part V: Research Frontiers.** In the end of the tutorial, we will briefly mention some of the recent research frontiers in recommendation community.

1. **Cross-domain recommendation.** Cross-domain or platform recommendation [28, 12], such as recommending a tweet to a Weibo user in China [12], is helpful to bridge the cultural gap caused by different platforms. The main challenge is how to address the more severe data sparsity issue, where the two platforms may have little content overlap.
2. **Reciprocal recommendation system.** In classical recommendation, we only care whether a user likes an item or not. But in some applications, such as online dating [32] and job recommendation, there should be a mutual interest between both parties when the recommendation is meaningful.
3. **Network embedding and text embedding for recommendation.** A recent trend [9, 33, 29, 3] is to learn a good representation for both entities and text for better recommendation. All sorts of network and text embedding techniques are emerging rapidly.

# 4. CONCLUSION

We systematically organize the cutting-edge recommendation algorithms that benefit from information network mining approaches into this tutorial, which turn out to be extremely effective in overcoming the sparsity and cold start issue in traditional recommendation methodologies. A broad spectrum of applications will be mentioned in this tutorial, which will benefit a wider audience.

## 5. ACKNOWLEDGMENTS

This work is partially supported by NSF Career #1453800.

## 6. REFERENCES


[1] M. Balabanovi, and Y. Shoham. Fab: content-based, collaborative recommendation. *Communications of the ACM*, 49(3):66–72, 1997.

[2] J. Chang and D. M. Blei. Relational topic models for document networks. In *AISTATS*, 2009.

[3] T. Chen and Y. Sun. Task-Guided and Path-Augmented Heterogeneous Network Embedding for Author Identification. In *WSDM*, 2017.

[4] G. Fu, Y. Ding, A. Seal, B. Chen, Y. Sun, and E. Bolton. Predicting drug target interactions using meta-path-based semantic network analysis In *BMC Bioinformatics 17:160*, 2016.

[5] H. Gao, J. Tang, and H. Liu. Exploring social-historical ties on location-based social networks. In *ICWSM*, 2012.

[6] H. Gao, J. Tang, X. Hu, and H. Liu. Exploring temporal effects for location recommendation on location-based social networks. In *RecSys*, pages 93–100, 2013.

[7] H. Gao, J. Tang, and H. Liu. Personalized location recommendation on location-based social networks. In *RecSys'14*, 2014.

[8] N. Gupta and S. Singh. Collectively embedding multi-relational data for predicting user preferences. *CoRR*, 2015.

[9] Y. Gu., B. Zhao, D. Hardtke, Y. Sun. Learning Global Term Weights for Content-based Recommender Systems. In *WWW*, 2016.

[10] B. Hu and M. Ester. Spatial topic modeling in online social media for location recommendation. In *RecSys*, pages 25–32, 2013.

[11] B. Liu, Y. Fu, Z. Yao, and H. Xiong. Learning geographical preferences for point-of-interest recommendation. In *KDD*, pages 1043–1051, 2013.

[12] X. Liu, T. Xia, Y. Yu, C. Guo, and Y. Sun. Cross Social Media Recommendation. In *ICWSM*, 2016.

[13] X. Liu, Y. Yu, C. Guo, and Y. Sun. Meta-Path-Based Ranking with Pseudo Relevance Feedback on Heterogeneous Graph for Citation Recommendation. In *CIKM*, 2014.

[14] X. Liu, Y. Yu, C. Guo, Y. Sun, L. Gao. Full-Text based Context-Rich Heterogeneous Network Mining Approach for Citation Recommendation. In *JCDL*, 2014.

[15] P. Lops, M. De Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. *Recommender systems handbook*, 73–105, Springer, 2011.

[16] Hao Ma, Michael R. Lyu, Irwin King. Learning to Recommend with Trust and Distrust Relationships. In *RecSys*, 2009.

[17] Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu, Irwin King. Recommender Systems with Social Regularization. In *WSDM*, 2011.

[18] Hao Ma. An Experimental Study on Implicit Social Recommendation. In *SIGIR*, 2013.

[19] R. M. Nallapati, A. Ahmed, E. P. Xing, and W. W. Cohen. Joint latent topic models for text and citations. In *KDD*, 2008.

[20] X. Ren, J. Liu, X. Yu, U. Khandelwal, Q. Gu, L. Wang, and J. Han. ClusCite: Effective Citation Recommendation by Information Network-Based Clustering. In *KDD*, 2014.

[21] X. Ren, A. El-Kishky, C. Wang, F. Tao, C. R. Voss, and J. Han. Clustype: Effective entity recognition and typing by relation phrase-based clustering. In *SIGKDD*, 2015.

[22] X. Ren, W. He, A. El-Kishky, H. Ji, C. R. Voss, M. Qu, and J. Han. Entity Typing: A Critical Step for Mining Structures from Massive Unstructured Text. In *MLG*, 2016.

[23] X. Ren, W. He, M. Qu, H. Ji, C. R. Voss, and J. Han. Label Noise Reduction in Entity Typing by Heterogeneous Partial-Label Embedding. In *SIGKDD*, 2016.

[24] T. Suddiqui, X. Ren, A. Parameswaran, and J. Han. FacetGist: Collective Extraction of Document Facets in Large Technical Corpora. In *CIKM*, 2016.

[25] Y. Sun, J. Han, X. Yan, P. S. Yu, and T. Wu. PathSim: Meta path-based top-k similarity search in heterogeneous information networks. In *VLDB'11/PVLDB*, Seattle, WA, Aug. 2011.

[26] Y. Sun, R. Barber, M. Gupta, C. C. Aggarwal, and J. Han, Co-Author Relationship Prediction in Heterogeneous Bibliographic Networks. In *ASONAM*, 2011.

[27] B. Wang, M. Ester, J. Bu, and D. Cai, Who Also Likes It? Generating the Most Persuasive Social Explanations in Recommender Systems In *AAAI*, 2014.

[28] Beidou Wang, Martin Ester, Yikang Liao, Jiajun Bu, Yu Zhu, Ziyu Guan, and Deng Cai, The Million Domain Challenge: Broadcast Email Prioritization by Cross-domain Recommendation, In *KDD*, 2016

[29] H. Wang, N. Wang, and D. Y. Yeung. Collaborative deep learning for recommender systems. In *KDD*, 2015.

[30] W. Wang, H. Yin, L. Chen, Y. Sun, S. Sadiq, and X. Zhou. Geo-sage: A geographical sparse additive generative model for spatial item recommendation. In *KDD*, 2015.

[31] W. Wang, H. Yin, S. Sadiq, L. Chen, M. Xie, and X. Zhou. Spore: A sequential personalized spatial item recommender system. In *ICDE*, 2016.

[32] P. Xia, B. Liu, Y. Sun, and C. X. Chen. Reciprocal Recommendation System for Online Dating. In *ASONAM*, 2015

[33] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, and S. Wang. Learning graph-based poi embedding for location-based recommendation. In *CIKM*, 2016.

[34] H. Yin and B. Cui. *Spatio-Temporal Recommendation in Social Media*. Springer Publishing Company, Incorporated, 1st edition, 2016.

[35] H. Yin, B. Cui, L. Chen, Z. Hu, and X. Zhou. Dynamic user modeling in social media systems. *ACM TOIS*, 33(3):10:1–10:44, Mar. 2015.

[36] H. Yin, B. Cui, Y. Sun, Z. Hu, and L. Chen. Lcars: A spatial item recommender system. *ACM TOIS*, 32(3):11:1–11:37, July 2014.

[37] H. Yin, Y. Sun, B. Cui, Z. Hu, and L. Chen. Lcars: A location-content-aware recommender system. In *KDD*, 2013.

[38] H. Yin, X. Zhou, B. Cui, H. Wang, K. Zheng, and Q. V. H. Nguyen. Adapting to user interest drift for poi recommendation. *IEEE Transactions on Knowledge and Data Engineering*, PP(99):1–14, June 2016.

[39] H. Yin, X. Zhou, Y. Shao, H. Wang, and S. Sadiq. Joint modeling of user check-in behaviors for point-of-interest recommendation. In *CIKM*, 2015.

[40] X. Yu, Q. Gu, M. Zhou, and J. Han. Citation Prediction in Heterogeneous Bibliographic Networks. In *SDM*, 2012.

[41] Xiao Yu, Hao Ma, Paul Hsu and Jiawei Han. On Building Entity Recommender Systems Using User Click Log and Freebase Knowledge. In *WSDM*, 2014.

[42] X. Yu, X. Ren, Y. Sun, B. Sturt, U. Khandelwal, Q. Gu, B. Norick, and J. Han. Personalized Entity Recommendation: A Heterogeneous Information Network Approach. In *WSDM*, 2014.

[43] X. Yu, X. Ren, Q. Gu, Y. Sun, and J. Han. Collaborative Filtering with Entity Similarity Regularization in Heterogeneous Information Networks. In *IJCAI-HINA*, 2013.

[44] X. Yu, X. Ren, Y. Sun, B. Sturt, U. Khandelwal, Q. Gu, B. Norick, and J. Han. HeteRec: Entity Recommendation in Heterogeneous Information Networks with Implicit User Feedback. In *RecSys*, 2013.

[45] Y. Zhu, X. Yan, L. Getoor, and C. Moore. Scalable text and link analysis with mixed-topic link models. In *KDD*, 2013.

[46] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*, pages 325–334, 2011.

[47] A. P. Singh and G. J. Gordon. Relational learning via collective matrix factorization. In *KDD*, pages 650–658, 2008.