Mining User Interests from Social Media

ABSTRACT

The abundance of user generated content on social media provides the opportunity to build models that are able to accurately and effectively extract, mine and predict users' interests with the hopes of enabling more effective user engagement, better quality delivery of appropriate services and higher user satisfaction. While traditional methods for building user profiles relied on AI-based preference elicitation techniques that could have been considered to be intrusive and undesirable by the users, more recent advances are focused on a non-intrusive yet accurate way of determining users' interests and preferences. In this tutorial, we will cover five important aspects related to the effective mining of user interests: we will introduce (1) the information sources that are used for extracting user interests, (2) the variety of types of user interest profiles that have been proposed in the literature, (3) techniques that have been adopted or proposed for mining user interests, (4) the scalability and resource requirements of the state of the art methods and, finally (5) the evaluation methodologies that are adopted in the literature for validating the appropriateness of the mined user interest profiles. We will also introduce existing challenges, open research questions and exciting opportunities for further work.

1 MOTIVATION AND OVERVIEW

Mining user interests from user behavioral data is critical for applications such as online advertising. Based on user interests, service providers such as advertisers, can significantly reduce service delivery costs by offering the most relevant products (e.g., ads) to their customers. The challenge of accurately and efficiently identifying user interests has been the subject of increasing attention in the past several years. Early approaches were based on explicit input from individuals about their own interests. To avoid the extra burden of manually filling in and maintaining interest profiles, most methods in the past two decades have focused on the development of techniques that can automatically and unobtrusively determine users' interests based on user behavioral data from data sources such as browsing history, page visits, the links they click on, the searches they perform and the topics they interact with [6, 7, 20, 22, 28].

With the emergence and growing popularity of social media such as blogging systems, wikis, social bookmarking, and microblogging services, many users are extensively engaged in at least some of these applications to express their feelings and views about a wide variety of social events/topics as they happen in real time by commenting, tagging, joining, sharing, liking, and publishing posts [1, 10, 29]. This has made social media an exciting and unique source of information about users' interests. For instance, when looking at Twitter data during the first week of March 2019, the rivalry between the two English Premier League soccer clubs, Tottenham Hotspur and Arsenal, is a topic that has attracted a lot of discussion and interest. The development of techniques that can automatically detect such topics and model users' interests towards them from online social media would be highly important and have the potential to improve the quality of applications that work on a user modeling basis, such as filtering Twitter streams [26], news recommendation [1] and retweet prediction [18], among others.

In this tutorial, we comprehensively introduce different strategies proposed in the literature, including our own work [4, 17, 36–39, 43, 44, 50, 51, 53–56], for mining user interests from social media with respect to the following five perspectives:

(1) *Information Sources*: The type of information sources used for extracting user interests from within social media such as textual content (comments, #tags), social network structure, and images [4, 12, 29, 45, 47]. Additionally, we will review external background knowledge sources such as semantic web resources and knowledge graphs that have been incorporated by some researchers to enhance the accuracy of user profiles [8, 25, 31, 51].

(2) *Profile Types:* Most of works in user interest mining from social media extract users' explicit interests that are directly observable from user content [2, 13, 30, 31, 48, 53]. However, given the increasingly noticeable free-rider, some other techniques focus on passive users and extract their implicit interests by considering the interaction patterns between users and topics [38, 45, 46, 50]. There is another line of work that is dedicated to predicting users' future interests instead of modeling current or past interests of users [5, 11, 24, 52]. These works are primarily focused on predicting if and which users would be interested in future topics on social media. The accurate identification of users' future interests on social media allows one to perform future planning by studying how users will react if certain topics emerge in the future.

(3) *Underlying Techniques*: Previous methods have employed different techniques to build user profiles including neural embeddings [16, 21, 24], collaborative filtering [3, 5, 14, 23, 33], topic modeling [23, 48, 55], link prediction [9, 32, 50, 55], regression [4, 19, 49], graph-based methods [15, 27, 41, 45, 53] and Semantic Web technologies [17, 25, 31, 35, 52]. We will review the techniques that have been used for identifying user interests and their different architectural variations.

(4) *Scalability and Resource Requirements:* Scalability is fundamental to user interest mining in order to accommodate torrents of social content. To this end, we provide a comprehensive overview of the speed-accuracy (efficiency-accuracy) trade-off when building user interest profiles for existing techniques of the literature [42]. In particular, we present a critical review of those which scale to online vs. offline for massive streaming social content.

(5) *Evaluation Methodology:* Intrinsic vs. extrinsic evaluations are two main evaluation techniques, which have been widely adopted in the literature. Intrinsic evaluation helps to assess the quality of the constructed user interest profiles based on user studies [8, 25, 34] while extrinsic evaluations measure the quality of the user interest profiles by looking at its impact on the effectiveness of other applications such as news recommendation and retweet prediction [1, 45, 53, 55]. We will review how each of these evaluation methodologies has been used in the literature.

2 TARGET AUDIENCE AND PREREQUISITES

The target audience for this tutorial will be those who have familiarity with social media mining and basics of data mining techniques. Where appropriate the tutorial will not make any assumptions about the audience's knowledge on more advanced techniques such as link prediction, matrix factorization, deep matching, entity linking and knowledge-graph based reasoning, among others. As such, sufficient details will be provided as appropriate so that the content will be accessible and understandable to those who have fundamental understanding of data mining principals. The tutorial will only assume familiarity with topics included in an undergraduate data mining course.

3 TUTORS

- Fattane Zarrinkalam, Ryerson university, *fzarrinkalam@ryerson.ca*
- Guangyuan Piao, NOKIA Bell Labs, *guangyuan.piao@nokia-bell-lab.com*
- Stefano Faralli, University of Rome Unitelma Sapienza, *stefano.faralli@unitelmasapienza.it*
- Ebrahim Bagheri, Ryerson university, *bagheri@ryerson.ca*

4 TUTORS BIO

Fattane Zarrinkalam is currently a Postdoctoral Fellow at the Laboratory of Systems, Software and Semantics (LS^3) at Ryerson University, where she works on projects related to Semantic-enabled Social Media Analysis. During her PhD studies, she focused on the identification of social media users' interests based on their individual and collective behavior on social media especially Twitter. She has already published her work in leading venues such as CIKM, ESWC and ECIR. In addition, she has published journal papers in some of the field's premier journals including *Information Retrieval* and *Information Processing and Management*. Further, during her PhD, she was involved in two patent applications that were filed with USPTO. She has presented a tutorial on User Interest Mining from Social Media at SIGIR2019 and KDD2019.

Guangyuan Piao is currently a research scientist at NOKIA Bell Labs with the research interest of AI for a wide range of applications. He did his PhD on Semantics-Aware User Modeling and Recommender Systems at the Insight Research Centre for Data Analytics (formerly DERI), NUI Galway, with the scholarship from SAP and SFI (Science Foundation Ireland). The research results have been published and presented in highly relevant venues such as UMUAI, CIKM, ECIR and Hypertext, and have been cited to pretentious venues. During his study, he worked as a teaching assistant in several courses related to Semantic Web and Object-oriented Programming.

Stefano Faralli is currently an Assistant Professor at University of Rome Unitelma Sapienza, with the research interest in recommender systems, social media user interest profiling, ontology learning, distributional semantics, word sense disambiguation/induction and Linked Open Data. He did his PhD on Computer Science at the University of Rome Sapienza and he was a postdoctoral researcher at Data and Web Science Group University of Mannheim. The research results have been published and presented in highly relevant venues such as IJCAI, ISWC, JWS, ICDM and JSNAM. He is currently a teacher of computer science courses (including Text Analytics and Social Network Analysis) and assistant teacher of Machine learning courses. He co-organized the International Social Interaction-based Recommendation (SIR 2018@CIKM).

Ebrahim Bagheri is an Associate Professor and the Director for the Laboratory for Systems, Software and Semantics (LS^3) at Ryerson University in Toronto. He also holds a Canada Research Chair (Tier II) in Software and Semantic Computing as well as an NSERC Industrial Research Chair in Social Media Analytics. He has been PI on projects worth over $8M funded by partners such as NSERC, AIF and IBM. Most recently in 2018, he was the Program Committee co-Chair for the Canadian Conference on Artificial Intelligence and also the Industry Program Committee co-Chair at IEEE/ACM International Conference on Advances in social media Analysis and Mining (ASONAM) and an Area Chair for NAACL-HLT 2019. He also serves on the Program Committee of venues such as RecSys and ICWSM as well guest-editor for international journals such as Information Systems and Information Processing and Management.

Ebrahim Bagheri's Ted talk on health research and its relation to social networks. Video is available here: https://www.youtube.com/watch?v=kS_TjzOQtNs

5 CORRESPONDING TUTOR

Ebrahim Bagheri: *bagheri@ryerson.ca*

6 TUTORIAL OUTLINE

This tutorial presents a comprehensive survey of user interest mining from online social media. In particular, this tutorial covers the following sections:

Session A [30 Minutes]: Background and Introduction to Theory of User Interest Mining The tutorial begins with a session about basics of user interest mining and various online social media. This includes preliminaries, motivations, and highlights on research questions to which user interest mining from online social media would provide an answer for. Then, we introduce different third-party applications that can take advantage of user interest mining from social network to improve the accuracy of their results. Finally, we review topics that will be covered in the tutorial followed by a disclaimer, i.e., what the tutorial is *not* about.

Session B [120 Minutes]: Techniques and Methods in User Interest Mining from Social Media Depending on the desirable type of user interest profiles, i.e., explicit or implicit or future user interest profiles, previous work have adopted different approaches for addressing the problem. Within these three categories, we lay out the details and provide a comparative analysis of existing methods in terms of their representation power, flexibility, resource needs and scalability. Specifically, in this session, we elaborate on how previous studies have used different techniques such as collaborative filtering [3, 5, 14, 23, 33], topic modeling [23, 40, 48, 55**?**], link prediction [9, 32, 50, 55], graph-based methods [15, 27, 41, 45, 53] and Semantic Web technologies [17, 25, 31, 35, 52] to construct a given type of interest profile for users (e.g. implicit interest profile).

Session C [30 Minutes]: Evaluation Methodologies, Future Directions and Open Challenges In this session, we first elaborate on different resources and two main approaches used in the literature to evaluate user interest profiles, namely intrinsic vs extrinsic evaluation techniques. Next, this session will present exciting open research questions in the state-of-the-art for mining users' interests from online social media. Accurate information extraction from online social media poses unique challenges due to the special characteristics of them. Social posts are rather short, noisy and informal and they often do not provide sufficient contextual information for identifying their semantics. In other words, the semantics of the context of the communicated information within a post is often implicit. Moreover, as a large majority of social network users are free-riders and cold start users, the interests of such users is challenging and they cannot be directly identified from their explicit contributions to the online social network. This tutorial presents the open issues that are important but have not been well addressed in recent studies which can inspire future directions in this research field. We will cover potential resources (e.g., Linked Open Data) and techniques (e.g. Learning-to-Rank, deep learning architectures and causal inference) that can be relevant for mining user interests.

7 RELEVANCE TO ARTIFICIAL INTELLIGENCE COMMUNITY AND A LIST OF RELATED TUTORIALS

The identification of user interests from social media has traditionally been of interest to the data mining community for at least three main reasons:

(1) Mining user interests targets the systematic extraction of information about users based on the behavioral signals, social network relations and users' content. The techniques that are used for this purpose are those that have been widely adopted and used in many aspects of data mining such as learning to rank, link prediction, graph mining, knowledge graphs, causal and predictive models, just to name a few.
(2) The accurate and complete detection of user interests finds relevance and importance in downstream data mining applications that rely on such information for decision making. For instance, effective recommendation requires sufficient context from the user who will be receiving the recommendations. Such context can be derived from users' interests that are mined from users' behavioral and interest profile.
(3) While traditional forms of user interest mining relied on techniques such as preference elicitation that could have been considered to be intrusive in many application areas, current methods for mining user interest profiles are based on mining publicly available online content. This provides an exciting opportunity to build large-scale, non-intrusive, scalable and efficient methods with large amounts of public data. As we will discuss in this tutorial, there are both intrinsic and extrinsic ways to effectively evaluate and benchmark the techniques in this area, which allow for reproducible and incremental studies.

It is important to note that while there has been similar synergistic tutorials on similar topics to this in other venues, the topic proposed in this tutorial distinguishes itself by focusing on ways to extract, mine and predict "user" level interest information. The following tutorials can be considered complementary and synergistic to the theme of our proposed tutorial:

(1) *User Group Analytics: Discovery, Exploration and Visualization* by Behrooz Omidvar-Tehrani and Sihem Amer-Yahia at CIKM 2018. This tutorial focuses on group level analytics while ours is focused on the user itself.
(2) *From Design to Analysis: Conducting Controlled Laboratory Experiments* with Users by Diane Kelly and Anita Crescenzi at SIGIR 2017. This tutorial is especially relevant to how experiments are designed for evaluating user interest mining techniques.
(3) *Knowledge Extraction and Inference from Text* by Soumen Chakrabarti at SIGIR 2018 (also CIKM 2017). Similar to the theme of our tutorial, this tutorial focused on extracting actionable knowledge from text. Our tutorial will go beyond text and cover other information types such as social relations and temporal characteristics.
(4) *Network Science of Teams: Characterization, Prediction, and Optimization* by Liangyue Li and Hanghang Tong at WSDM 2018. Similar to User Group Analytics tutorial presented above, this tutorial focuses on team level dynamics whereas our tutorial will be focused on user-level analytics and hence are complementary.
(5) *Behavior Analytics: Methods and Applica* by Longbing Cao, Philip S Yu and Guansong Pang at KDD 2018. This tutorial focuses on behavior analytics of customers at group level analytics while our tutorial is focused on the user itself.
(6) *Social Media Analytics: Tracking, Modeling and Predicting the Flow of Information through Networks* by Jure Leskovec at KDD 2011. This tutorial focused on user interactions in social media to track the flow of relevant information and predict missing links. Our tutorial utilizes both user content and interactions to more specifically extract users' interests from social media.

There are many other similar tutorials presented at major venues similar to the above. We aim to provide a complementary view of analytics at the user-level, while past tutorials focused on other aspects such as group or team-level analytics.

REFERENCES

[1] Fabian Abel, Qi Gao, Geert-Jan Houben, and Ke Tao. 2011. Analyzing User Modeling on Twitter for Personalized News Recommendations. In *User Modeling, Adaptation and Personalization - 19th International Conference, UMAP 2011, Girona, Spain, July 11-15, 2011. Proceedings*. 1–12. https://doi.org/10.1007/978-3-642-22362-4_1
[2] Fabian Abel, Qi Gao, Geert-Jan Houben, and Ke Tao. 2011. Semantic Enrichment of Twitter Posts for User Profile Construction on the Social Web. In *The Semantic Web: Research and Applications - 8th Extended Semantic Web Conference, ESWC 2011, Heraklion, Crete, Greece, May 29 - June 2, 2011, Proceedings, Part II*. 375–389. https://doi.org/10.1007/978-3-642-21064-8_26
[3] Amr Ahmed, Bhargav Kanagal, Sandeep Pandey, Vanja Josifovski, Lluis Garcia Pueyo, and Jeffrey Yuan. 2013. Latent factor models with additive and hierarchically-smoothed user preferences. In *Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy, February 4-8, 2013*. 385–394. https://doi.org/10.1145/2433396.2433445
[4] Negar Arabzadeh, Hossein Fani, Fattane Zarrinkalam, Ahmed Navivala, and Ebrahim Bagheri. 2018. Causal Dependencies for Future Interest Prediction on Twitter. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22-26, 2018*. 1511–1514. https://doi.org/10.1145/3269206.3269312
[5] Hongyun Bao, Qiudan Li, Stephen Shaoyi Liao, Shuangyong Song, and Heng Gao. 2013. A new temporal and social PMF-based method to predict users' interests in micro-blogging. *Decision Support Systems* 55, 3 (2013), 698–709. https://doi.org/10.1016/j.dss.2013.02.007
[6] Alex Beutel, Leman Akoglu, and Christos Faloutsos. 2015. Fraud Detection through Graph-Based User Behavior Modeling. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, October 12-16, 2015*. 1696–1697. https://doi.org/10.1145/2810103.2812702

[7] Alex Beutel, Leman Akoglu, and Christos Faloutsos. 2015. Graph-Based User Behavior Modeling: From Prediction to Fraud Detection. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015.* 2309–2310. https://doi.org/10.1145/2783258.2789985

[8] C. Budak, A. Kannan, R. Agrawal, , and J. Pedersen. 2014. Inferring User Interests From Microblogs. In *Technical Report, MSR-TR-2014-68.*

[9] Charalampos Chelmis and Viktor K. Prasanna. 2013. Social Link Prediction in Online Social Tagging Systems. *ACM Trans. Inf. Syst.* 31, 4 (2013), 20:1–20:27. https://doi.org/10.1145/2516891

[10] Pin-Yu Chen, Chun-Chen Tu, Pai-Shun Ting, Ya-Yun Lo, Danai Koutra, and Alfred O. Hero III. 2016. Identifying Influential Links for Event Propagation on Twitter: A Network of Networks Approach. *CoRR* abs/1609.05378 (2016). arXiv:1609.05378 http://arxiv.org/abs/1609.05378

[11] Wei Chen, Wynne Hsu, and Mong-Li Lee. 2013. Modeling user's receptiveness over time for recommendation. In *The 36th International ACM SIGIR conference on research and development in Information Retrieval, SIGIR '13, Dublin, Ireland - July 28 - August 01, 2013.* 373–382.

[12] Pravallika Devineni, Danai Koutra, Michalis Faloutsos, and Christos Faloutsos. 2017. Facebook wall posts: a model of user behaviors. *Social Netw. Analys. Mining* 7, 1 (2017), 6:1–6:15. https://doi.org/10.1007/s13278-017-0422-9

[13] Pravallika Devineni, Evangelos E. Papalexakis, Danai Koutra, A. Seza Doğruöz, and Michalis Faloutsos. 2017. One Size Does Not Fit All: Profiling Personalized Time-Evolving User Behaviors. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, Sydney, Australia, July 31 - August 03, 2017.* 331–340. https://doi.org/10.1145/3110025.3110050

[14] Ernesto Diaz-Aviles, Lucas Drumond, Zeno Gantner, Lars Schmidt-Thieme, and Wolfgang Nejdl. 2012. What is happening right now ... that interests me?: online topic discovery and recommendation in twitter. In *21st ACM International Conference on Information and Knowledge Management, CIKM'12, Maui, HI, USA, October 29 - November 02, 2012.* 1592–1596. https://doi.org/10.1145/2396761.2398479

[15] Yuxin Ding, Shengli Yan, Yibin Zhang, Wei Dai, and Li Dong. 2016. Predicting the attributes of social network users using a graph-based machine learning method. *Computer Communications* 73 (2016), 3–11. https://doi.org/10.1016/j.comcom.2015.07.007

[16] Hossein Fani, Ebrahim Bagheri, and Weichang Du. 2017. Temporally Like-minded User Community Identification through Neural Embeddings. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management.* 577–586. https://doi.org/10.1145/3132847.3132955

[17] Stefano Faralli, Giovanni Stilo, and Paola Velardi. 2017. Automatic acquisition of a taxonomy of microblogs users' interests. *J. Web Semant.* 45 (2017), 23–40. https://doi.org/10.1016/j.websem.2017.05.004

[18] Wei Feng and Jianyong Wang. 2013. Retweet or not?: personalized tweet reranking. In *Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy, February 4-8, 2013.* 577–586. https://doi.org/10.1145/2433396.2433470

[19] Li Gao, Jia Wu, Chuan Zhou, and Yue Hu. 2017. Collaborative Dynamic Sparse Topic Regression with User Profile Evolution for Item Recommendation. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA.* 1316–1322. http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14438

[20] Fabio Gasparetti. 2017. Modeling user interests from web browsing activities. *Data Min. Knowl. Discov.* 31, 2 (2017), 502–547. https://doi.org/10.1007/s10618-016-0482-x

[21] Sadid A. Hasan, Yuan Ling, Joey Liu, and Oladimeji Farri. 2015. Exploiting Neural Embeddings for Social Media Data Analysis. In *Proceedings of The Twenty-Fourth Text REtrieval Conference, TREC 2015, Gaithersburg, Maryland, USA, November 17-20, 2015.*

[22] Michal Holub and Mária Bieliková. 2010. Estimation of user interest in visited web page. In *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010.* 1111–1112. https://doi.org/10.1145/1772690.1772829

[23] Liangjie Hong, Aziz S. Doumith, and Brian D. Davison. 2013. Co-factorization machines: modeling user interests and predicting individual decisions in Twitter. In *Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy, February 4-8, 2013.* 557–566. https://doi.org/10.1145/2433396.2433467

[24] Jaeyong Kang, Hongseok Choi, and Hyunju Lee. 2019. Deep recurrent convolutional networks for inferring user interests from social media. *J. Intell. Inf. Syst.* 52, 1 (2019), 191–209.

[25] Pavan Kapanipathi, Prateek Jain, Chitra Venkatramani, and Amit P. Sheth. 2014. User Interests Identification on Twitter Using a Hierarchical Knowledge Base. In *The Semantic Web: Trends and Challenges - 11th International Conference, ESWC 2014, Anissaras, Crete, Greece, May 25-29, 2014. Proceedings.* 99–113. https://doi.org/10.1007/978-3-319-07443-6_8

[26] Pavan Kapanipathi, Fabrizio Orlandi, Amit P. Sheth, and Alexandre Passant. 2011. Personalized Filtering of the Twitter Stream. In *Proceedings of the second Workshop on Semantic Personalized Information Management: Retrieval and Recommendation 2011, Bonn, Germany, October 24, 2011.* 6–13.

[27] Aparna Krishnan and Raghav Ramesh. [n. d.]. Graph based User Interest Modeling in Twitter CS 224W: Final Project, Group 41. ([n. d.]).

[28] Jun Li and Peng Zhang. 2013. Mining Explainable User Interests from Scalable User Behavior Data. In *Proceedings of the First International Conference on Information Technology and Quantitative Management.* 789–796. https://doi.org/10.1016/j.procs.2013.05.101

[29] Xin Li, Lei Guo, and Yihong Eric Zhao. 2008. Tag-based social interest discovery. In *Proceedings of the 17th International Conference on World Wide Web, WWW 2008, Beijing, China, April 21-25, 2008.* 675–684. https://doi.org/10.1145/1367497.1367589

[30] Shangsong Liang, Zhaochun Ren, Yukun Zhao, Jun Ma, Emine Yilmaz, and Maarten de Rijke. 2017. Inferring Dynamic User Interests in Streams of Short Texts for User Clustering. *ACM Trans. Inf. Syst.* 36, 1 (2017), 10:1–10:37. https://doi.org/10.1145/3072606

[31] Matthew Michelson and Sofus A. Macskassy. 2010. Discovering users' topics of interest on twitter: a first look. In *Proceedings of the Fourth Workshop on Analytics for Noisy Unstructured Text Data, AND 2010, Toronto, Ontario, Canada, October 26th, 2010 (in conjunction with CIKM 2010).* 73–80. https://doi.org/10.1145/1871840.1871852

[32] Alan Mislove, Bimal Viswanath, P. Krishna Gummadi, and Peter Druschel. 2010. You are who you know: inferring user profiles in online social networks. In *Proceedings of the Third International Conference on Web Search and Web Data Mining, WSDM 2010, New York, NY, USA, February 4-6, 2010.* 251–260. https://doi.org/10.1145/1718487.1718519

[33] Makoto Nakatsuji, Yasuhiro Fujiwara, Toshio Uchiyama, and Hiroyuki Toda. 2012. Collaborative Filtering by Analyzing Dynamic User Interests Modeled by Taxonomy. In *The Semantic Web - ISWC 2012 - 11th International Semantic Web Conference, Boston, MA, USA, November 11-15, 2012, Proceedings, Part I.* 361–377. https://doi.org/10.1007/978-3-642-35176-1_23

[34] Fedelucio Narducci, Cataldo Musto, Giovanni Semeraro, Pasquale Lops, and Marco de Gemmis. 2013. Leveraging Encyclopedic Knowledge for Transparent and Serendipitous User Profiles. In *User Modeling, Adaptation, and Personalization - 21th International Conference, UMAP 2013, Rome, Italy, June 10-14, 2013, Proceedings.* 350–352. https://doi.org/10.1007/978-3-642-38844-6_36

[35] Fabrizio Orlandi, John G. Breslin, and Alexandre Passant. 2012. Aggregated, interoperable and multi-domain user profiles for the social web. In *I-SEMANTICS 2012 - 8th International Conference on Semantic Systems, I-SEMANTICS '12, Graz, Austria, September 5-7, 2012.* 41–48. https://doi.org/10.1145/2362499.2362506

[36] Guangyuan Piao and John G. Breslin. 2016. Analyzing Aggregated Semantics-enabled User Modeling on Google+ and Twitter for Personalized Link Recommendations. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP 2016, Halifax, NS, Canada, July 13 - 17, 2016.* 105–109. https://doi.org/10.1145/2930238.2930278

[37] Guangyuan Piao and John G. Breslin. 2016. User Modeling on Twitter with WordNet Synsets and DBpedia Concepts for Personalized Recommendations. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management, CIKM 2016, Indianapolis, IN, USA, October 24-28, 2016.* 2057–2060. https://doi.org/10.1145/2983323.2983908

[38] Guangyuan Piao and John G. Breslin. 2017. Inferring User Interests for Passive Users on Twitter by Leveraging Followee Biographies. In *ECIR.* 122–133.

[39] Guangyuan Piao and John G. Breslin. 2018. Inferring user interests in microblogging social networks: a survey. *User Model. User-Adapt. Interact.* 28, 3 (2018), 277–329. https://doi.org/10.1007/s11257-018-9207-8

[40] Daniel Ramage, Susan T. Dumais, and Daniel J. Liebling. 2010. Characterizing Microblogs with Topic Models. In *Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23-26, 2010.* http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1528

[41] Yongwook Shin, Chuh Yeop Ryo, and Jonghun Park. 2014. Automatic extraction of persistent topics from social text streams. *World Wide Web* 17, 6 (2014), 1395–1420. https://doi.org/10.1007/s11280-013-0251-3

[42] Nemanja Spasojevic, Jinyun Yan, Adithya Rao, and Prantik Bhattacharyya. 2014. LASTA: large scale topic assignment on multiple social networks. In *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, New York, NY, USA - August 24 - 27, 2014.* 1809–1818. https://doi.org/10.1145/2623330.2623350

[43] Giorgia Di Tommaso, Stefano Faralli, Giovanni Stilo, and Paola Velardi. 2018. Wiki-MID: A Very Large Multi-domain Interests Dataset of Twitter Users with Mappings to Wikipedia. In *17th International Semantic Web Conference.* 36–52.

[44] Anil Kumar Trikha, Fattane Zarrinkalam, and Ebrahim Bagheri. 2018. Topic-Association Mining for User Interest Detection. In *ECIR.* 665–671.

[45] Jinpeng Wang, Wayne Xin Zhao, Yulan He, and Xiaoming Li. 2014. Infer User Interests via Link Structure Regularization. *ACM TIST* 5, 2 (2014), 23:1–23:22. https://doi.org/10.1145/2499380

[46] Tingting Wang, Hongyan Liu, Jun He, and Xiaoyong Du. 2013. Mining User Interests from Information Sharing Behaviors in Social Media. In *Advances in Knowledge Discovery and Data Mining, 17th Pacific-Asia Conference, PAKDD*

2013, Gold Coast, Australia, April 14-17, 2013, Proceedings, Part II. 85–98. https://doi.org/10.1007/978-3-642-37456-2_8
[47] Pengtao Xie, Yulong Pei, Yuan Xie, and Eric P. Xing. 2015. Mining User Interests from Personal Photos. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.* 1896–1902. http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9655
[48] Zhiheng Xu, Rong Lu, Liang Xiang, and Qing Yang. 2011. Discovering User Interest on Twitter with a Modified Author-Topic Model. In *Proceedings of the 2011 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2011, Campus Scientifique de la Doua, Lyon, France, August 22-27, 2011.* 422–429. https://doi.org/10.1109/WI-IAT.2011.47
[49] Lei Yang, Tao Sun, Ming Zhang, and Qiaozhu Mei. 2012. We know what @you #tag: does the dual role affect hashtag adoption?. In *Proceedings of the 21st World Wide Web Conference 2012, WWW 2012, Lyon, France, April 16-20, 2012.* 261–270. https://doi.org/10.1145/2187836.2187872
[50] Fattane Zarrinkalam, Hossein Fani, Ebrahim Bagheri, and Mohsen Kahani. 2016. Inferring Implicit Topical Interests on Twitter. In *Advances in Information Retrieval - 38th European Conference on IR Research, ECIR 2016, Padua, Italy, March 20-23, 2016. Proceedings.* 479–491. https://doi.org/10.1007/978-3-319-30671-1_35
[51] Fattane Zarrinkalam, Hossein Fani, Ebrahim Bagheri, and Mohsen Kahani. 2017. Predicting Users' Future Interests on Twitter. In *Advances in Information Retrieval - 39th European Conference on IR Research, ECIR 2017, Aberdeen, UK, April 8-13, 2017, Proceedings.* 464–476.
[52] Fattane Zarrinkalam, Hossein Fani, Ebrahim Bagheri, and Mohsen Kahani. 2017. Predicting Users' Future Interests on Twitter. In *Advances in Information Retrieval - 39th European Conference on IR Research, ECIR 2017, Aberdeen, UK, April 8-13, 2017, Proceedings.* 464–476. https://doi.org/10.1007/978-3-319-56608-5_36
[53] Fattane Zarrinkalam, Hossein Fani, Ebrahim Bagheri, Mohsen Kahani, and Weichang Du. 2015. Semantics-Enabled User Interest Detection from Twitter. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015, Singapore, December 6-9, 2015 - Volume I.* 469–476. https://doi.org/10.1109/WI-IAT.2015.182
[54] Fattane Zarrinkalam, Stefano Faralli, Guangyuan Piao, and Ebrahim Bagheri. 2020. Extracting, Mining and Predicting Users' Interests from Social Media. *submited to Foundations and Trends® in Information Retrieval* (2020).
[55] Fattane Zarrinkalam, Mohsen Kahani, and Ebrahim Bagheri. 2018. Mining user interests over active topics on social networks. *Inf. Process. Manage.* 54, 2 (2018), 339–357. https://doi.org/10.1016/j.ipm.2017.12.003
[56] Fattane Zarrinkalam, Mohsen Kahani, and Ebrahim Bagheri. 2018. User interest prediction over future unobserved topics on social networks. *Information Retrieval Journal* (2018), 1–36.