Mobility and Social Networking: A Data Management Perspective *

Mohamed F. Mokbel
Dept. of Computer Science and Engineering
University of Minnesota, Twin Cities
Minneapolis, MN 55455
mokbel@cs.umn.edu

Mohamed Sarwat
Dept. Computer Science and Engineering
University of Minnesota, Twin Cities
Minneapolis, MN 55455
sarwat@cs.umn.edu

ABSTRACT

This tutorial presents the state-of-the-art research that lies at the intersection of two hot topics in the data management community: (1) *social networking* and (2) *mobility*. In this tutorial, we give an overview of existing research work, systems, and applications related to both social networking and mobility. In addition, we introduce several resources (i.e., datasets, software tools) as well as a list of promising research directions.

1. TUTORIAL OUTLINE

Online social networks, such as Facebook and Twitter have become very popular in the past decade. Users register to online social networks in order to keep in touch with their friends and family, learn about their news, get recommendations from them, and engage in online social events. As mobile devices (e.g., smart phones, GPS devices) became ubiquitous, location-based social networking services (e.g., Foursquare and Facebook Places) are getting more and more popular. For instance, as of September 2012, Foursquare claims to have over 25 million people worldwide, and over billions of check-ins with millions more every day. Users, in a location-based social network, are associated with a geo-location, and might alert friends when visiting a venue (e.g., restaurant, bar) by checking-in on their mobile phones (e.g., iPhone, Android). The rise of location-based social networking applications has led to the emergence of both social networking and mobility side by side, which led to the rise of new research challenges and opportunities. This tutorial presents the state-of-the-art research that lies at the intersection of both: Social Networking and Mobility.

The merger between social networking and mobility [33, 4, 8, 46] brought together the following data types: (1) *Social Networking data*: represents the friendship between different users (usually represented by a social graph) as well as all sorts of social interactions between users. (2) *Spatial/Spatio-temporal data*: represents the users geo-locations, venues (e.g., restaurant, gym, shopping mall) geo-locations and information about users visiting different places at different times. (3) *Users Opinions data*: represents how much a user likes the places she visits by expressing (e.g., Alice visited restaurant A and gave it a rating of five over five). We then illustrate how different mixes of this data trilogy has been leveraged to explore new trends and by developers to build novel mobile applications [4, 8, 16, 42].

Social networking data management research is mainly concerned with managing users social interactions and collaboration, storing / retrieving social media (e.g., Microblogs, News Feed), and analyzing users behavior. Mobile data management research focuses on handling user GeoSpatial location and contextual information. This tutorial takes an overarching approach by surveying the research that combines both social networking and mobility from the following perspectives: (1) Microblog search and social news feed queries, (2) Recommendation Services, (3) Analytics, (4) Crowdsourcing, (5) System and Media Visualization, and (6) Risks and Threats. In summary, the tutorial consists of the following parts:

1. *PART I: Microblog Search and Social News Feed Queries*: In this part, we first describe how both geo-location and social awareness work in concert to answer queries [3, 1, 7, 6, 5, 13, 21, 30, 35, 38, 41]. We present recent studies that show how both spatial and social aspects can be combined to enhance Microblog search and social news feed quality. We also analyze existing GeoSocial Microblog search and Location-aware New feed querying system from a scalability and efficiency perspective.
2. *PART II: Recommendation Services*: We present recent studies which show that geo-location matters in recommender systems, and we manifest several techniques to incorporate the spatial/spatio-temporal information [17, 34, 37, 39, 43] side-by-side with users opinions data in traditional recommender systems [16, 28, 36, 31]. We also highlight the research works that leverage GeoSocial data points and trajectories for travel and itinerary recommendations [47, 45].
3. *PART III: Analytics*: We give an overview of computational techniques that are harnessed to analyze GeoSocial data [9, 23, 20, 44] to learn more about human behavior, the societal and economical consequences of such analysis [19, 2].
4. *PART IV: Crowdsourcing*: We summarize research work performed in the Volunteered Geographic Information (VGI) and participatory sensing area. We then survey recent papers that address the crowdsourcing topic [15, 11, 24, 25, 27, 26] from a mobility perspective [10, 14].
5. *PART V: Systems and Media Visualization*: we highlight systems that aim at visualizing GeoSocial Media [22, 33, 48] (e.g., Geo-Tagged Tweets, Geo-Tagged Videos) and social

*This work is supported in part by the National Science Foundation under Grants IIS-0952977 and IIS-1218168.

Proceedings of the VLDB Endowment, Vol. 6, No. 11

interactions between mobile users. We present the state-of-the-art GeoSocial visualization techniques [32, 18].

6. *PART VI: Risks and Threats*: In this part, we throw spotlight on the main risks that may arise from combining social networking and mobility. For instance, we highlight the possible threats of revealing user location in online social networking websites. We then highlight recent research work that aims at preserving the user location and absence privacy in location-aware social networking services (e.g., privacy [29, 40, 12]).

For all aforementioned topics, we present results from recent research work, case studies featuring hot mobile applications, and the anatomy of built systems. Finally, we conclude the tutorial by summarizing and presenting open research directions.

2. REFERENCES

[1] J. Bao, M. F. Mokbel, and C.-Y. Chow. GeoFeed: A Location-Aware News Feed System. In *ICDE*, 2012.

[2] S. Bhagat, A. Goyal, and L. V. S. Lakshmanan. Maximizing product adoption in social networks. In *WSDM*, 2012.

[3] M. Busch, K. Gade, B. Larson, P. Lok, S. Luckenbill, and J. Lin. Earlybird: Real-Time Search at Twitter. In *ICDE*, 2012.

[4] Y. Cai and T. Xu. Design, analysis, and implementation of a large-scale real-time location-based information sharing system. In *MobiSys*, 2008.

[5] C. C. CAO, J. She, Y. Tong, and L. Chen. Whom to ask? jury selection for decision making tasks on micro-blog services. In *VLDB*, 2012.

[6] C. Chen, F. Li, B. C. Ooi, and S. Wu. Ti: an efficient indexing mechanism for real-time search on tweets. In *SIGMOD*, 2011.

[7] L. Chen, G. Cong, C. S. Jensen, and D. Wu. Spatial Keyword Query Processing: An Experimental Evaluation. In *VLDB*, 2013.

[8] Y. Chen, K. Jiang, Y. Zheng, C. Li, and N. Yu. Trajectory simplification method for location-based social networking services. In *GIS-LBSN*, 2009.

[9] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *KDD*, 2011.

[10] A. Efentakis, D. Theodorakis, and D. Pfoser. Crowdsourcing computing resources for shortest-path computation. In *GIS*, 2012.

[11] A. Feng, M. J. Franklin, D. Kossmann, T. Kraska, S. Madden, S. Ramesh, A. Wang, and R. Xin. CrowdDB: Query Processing with the VLDB Crowd. *PVLDB*, 2011.

[12] D. Freni, C. R. Vicente, S. Mascetti, C. Bettini, and C. S. Jensen. Preserving location and absence privacy in geo-social networks. In *CIKM*, 2010.

[13] N. Gupta, L. Kot, S. Roy, G. Bender, J. Gehrke, and C. Koch. Entangled queries: Enabling declarative data-driven coordination. *TODS*, 37(3):21, 2012.

[14] L. Kazemi and C. Shahabi. Geocrowd: Enabling query answering with spatial crowdsourcing. In *GIS*, 2012.

[15] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *CHI*, 2008.

[16] K. Kodama, Y. Iijima, X. Guo, and Y. Ishikawa. Skyline queries based on user locations and preferences for making location-based recommendations. In *GIS-LBSN*, 2009.

[17] J. J. Levandoski, M. Sarwat, A. Eldawy, and M. F. Mokbel. LARS: A Location-Aware Recommender System. In *ICDE*, 2012.

[18] M. D. Lieberman and H. Samet. Supporting rapid processing and interactive map-based exploration of streaming news. In *GIS*, 2012.

[19] W. Lu and L. V. S. Lakshmanan. Profit maximization over social networks. In *ICDM*, 2012.

[20] W. Lu, Y. Shen, S. Chen, and B. C. Ooi. Efficient processing of k nearest neighbor joins using mapreduce. In *VLDB*, 2012.

[21] K. Mamouras, S. Oren, L. Seeman, L. Kot, and J. Gehrke. The Complexity of Social Coordination. *PVLDB*, 5(11):1172–1183, 2012.

[22] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. Tweets as data: demonstration of TweeQL and Twitinfo. In *SIGMOD*, 2011.

[23] G. D. F. Morales, A. Gionis, and M. Sozio. Social content matching in mapreduce. In *VLDB*, 2011.

[24] A. G. Parameswaran, H. Garcia-Molina, H. Park, N. Polyzotis, A. Ramesh, and J. Widom. CrowdScreen: algorithms for filtering data with humans. In *SIGMOD*, pages 361–372, 2012.

[25] A. G. Parameswaran, H. Park, H. Garcia-Molina, N. Polyzotis, and J. Widom. Deco: declarative crowdsourcing. In *CIKM*, 2012.

[26] H. Park, R. Pang, A. G. Parameswaran, H. Garcia-Molina, N. Polyzotis, and J. Widom. An overview of the deco system: data model and query language; query processing and optimization. *SIGMOD Record*, 41(4):22–27, 2012.

[27] H. Park, R. Pang, A. G. Parameswaran, H. Garcia-Molina, N. Polyzotis, and J. Widom. Deco: A System for Declarative Crowdsourcing. *PVLDB*, 5(12):1990–1993, 2012.

[28] M.-H. Park et al. Location-based recommendation system using bayesian user's preference model in mobile devices. In *UIC*, 2007.

[29] K. P. N. Puttaswamy and B. Y. Zhao. Preserving privacy in location-based mobile social applications. In *HotMobile*, 2010.

[30] S. B. Roy and K. Chakrabarti. Location-aware type ahead search on spatial databases: semantics and efficiency. In *SIGMOD*, 2011.

[31] S. B. Roy et al. Space efficiency in group recommendation. *VLDB J.*, 19(6):877–900, 2010.

[32] A. D. Sarma, H. Lee, H. Gonzalez, J. Madhavan, and A. Y. Halevy. Efficient spatial sampling of large geographical tables. In *SIGMOD*, 2012.

[33] M. Sarwat, J. Bao, A. Eldawy, J. J. Levandoski, A. Magdy, and M. F. Mokbel. Sindbad: A Location-based Social Networking System. In *SIGMOD*, 2012.

[34] M. Sarwat, J. J. Levandoski, A. Eldawy, and M. F. Mokbel. LARS*: A Scalable and Efficient Location-Aware Recommender System. In *TKDE*, 2013.

[35] A. Silberstein, J. Terrace, B. F. Cooper, and R. Ramakrishnan. Feeding frenzy: Selectively materializing user's event feed. In *SIGMOD*, 2010.

[36] Y. Takeuchi and M. Sugimoto. An outdoor recommendation system based on user location history. In *UIC*, 2006.

[37] S. Thirumuruganathan, M. Das, S. Desai, S. Amer-Yahia, G. Das, and C. Yu. MapRat: Meaningful Explanation, Interactive Exploration and Geo-Visualization of Collaborative Ratings. *PVLDB*, 5(12):1986–1989, 2012.

[38] J. R. Thomsen, M. L. Yiu, and C. S. Jensen. Effective caching of shortest paths for location-based services. In *SIGMOD*, 2012.

[39] P. Venetis, H. Gonzalez, C. S. Jensen, and A. Y. Halevy. Hyper-local, directions-based ranking of places. *VLDB*, 4(5):290–301, 2011.

[40] W. Wei, F. Xu, and Q. Li. MobiShare: Flexible privacy-preserving location sharing in mobile online social networks. In *INFOCOM*, 2012.

[41] W. Xu, C.-Y. Chow, M. L. Yiu, Q. Li, and C. K. Poon. MobiFeed: a location-aware news feed system for mobile users. In *GIS*, 2012.

[42] M. Ye, P. Yin, and W.-C. Lee. Location recommendation for location-based social networks. In *GIS*, 2010.

[43] H. Yin, B. Cui, J. Li, J. Yao, and C. Chen. Challenging the Long Tail Recommendation. *PVLDB*, 5(9):896–907, 2012.

[44] C. Zhang, L. Shou, K. Chen, G. Chen, and Y. Bei. Evaluating geo-social influence in location-based social networks. In *CIKM*, 2012.

[45] V. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative Location and Activity Recommendations with GPS History Data. In *WWW*, 2010.

[46] Y. Zheng, Y. Chen, X. Xie, and W.-Y. Ma. Geolife2.0: A location-based social networking service. In *MDM*, 2009.

[47] Y. Zheng and X. Xie. Learning travel recommendations from user-generated GPS traces. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(1):2, 2011.

[48] Y. Zheng, X. Xie, and W.-Y. Ma. GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. *IEEE Data Eng. Bull.*, 33(2):32–39, 2010.