2016 IEEE 16th International Conference on Data Mining Workshops

Infer Mobility Patterns and Social Dynamics for Modelling Human Behaviour

Luca Luceri
SUPSI, Manno, Switzerland and University of Bern, Bern, Switzerland
Email: luca.luceri@supsi.ch

***Abstract*—Investigating human mobility patterns and comprehending the social dynamics that govern people movements is of high interest for multiple aspects and reasons. Location-based services, mobile network management, and urban planning are just few of the several applications that benefit from this kind of assessment. This work focuses on the stochastic analysis of spatiotemporal and social network data in order to build a human behaviour model which aims to predict social dynamics and to infer users' mobility patterns and interests.**

***Keywords*-Social Network Analysis (SNA); spatiotemporal data mining; mobility prediction; human behaviour model.**

I. Introduction and Motivation

The understanding of social dynamics, and in particular of the inter-dependence with people movements, has recently received attention both in academy and in industry. In fact, the ability to find patterns in human mobility and to predict next visited places represent a valuable input in numerous applications. These estimates play a crucial role in various fields, ranging from location-based services and urban planning to mobile network design and management.

Studies on mobility data (e.g., GPS traces, Call Detail Record (CDR)) showed that human movements exhibit a high regularity and recurrence, allowing a certain degree of predictability [1], [2]. A number of models have been proposed to reproduce mobility patterns and predict users' next visited locations. Order-k Markov, Bayesian network, Support Vector Machine (SVM), decision tree and several heuristic approaches are just some of the employed solutions utilized to accomplish these tasks. The major open point is how to enhance the accuracy of the inference. One of the envisioned solutions is to exploit additional information, e.g., users' social relationships and interests. Several approaches have been proposed to separately model social networks and mobility dynamics. However, only recently researchers started exploring the interplay between human movement and social ties of individuals, finding a high correlations between them [3], [4]. Results show that physical location and social ties are strongly interlaced with a two-fold effect: social relationships are influenced by physical location [5], and mobility prediction accuracy improves by utilizing location information related to users' friends and acquaintances [6]. Moreover, also mass phenomenon as segregation and gentrification combined with socioeconomic and demographic information can be conveniently utilized to understand large-scale human dynamics. However, few analyses have been carried out combining geo-social information on a large-scale [7], [8]. This altogether brings to the quest for general models that combine social networks and mobility dynamics data.

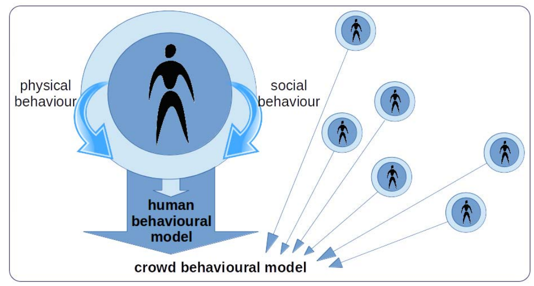

Figure 1. The two-level behavioural model

The purpose of this work is to utilize geospatial information along with data collected from social network platforms to produce an understanding and modelling of the human behaviour. This will be done both at user and at crowd level, and combining the two of them. This kind of analysis is of great interest not only to perform social studies and to enhance location-based services but can also be applied in some specific application scenario. As a potential application of the developed methodology we will consider the deployment and self-optimization of next generation (5G) telecommunication networks, with the purpose of improving network performances in terms of data rate, latency and resources management. We expect that this work might provide the modelling needed for understanding and forecasting telecommunication network conditions. In fact, the proposed model will allow to predict users' mobility and social dynamics, to estimate spatial and temporal density of subjects in a certain area and to assess the contents they will be interested in. These estimates will be beneficial for a self-management system in order to automatically and dynamically tune network key parameters according to the forecasted scenario.

II. Research Goals

This work aims to analyze and utilize users' social, physical and contextual information in order to build a *two-level behavioural model*. The two layers, sketched in Figure 1, concern the granularity of the study: it will be carried out both at single-user and at crowd level. Finally, these two

2375-9259/16 $31.00 © 2016 IEEE
DOI 10.1109/ICDMW.2016.187

layers will be coupled to combine and exploit both kinds of information, e.g., if there is not available information related to a user in a specific time we can utilize the crowd level information for the estimates. In order to achieve this aim the following steps are foreseen:

- *Spatiotemporal data mining:* analysis of users' geolocation data in order to extract spatial and time features to predict and model users' mobility. This will be performed both in terms of single-user and of groups or communities.
- *Social Network Analysis (SNA):* examine social relationships, study mass phenomenon and investigate the interdependence between physical location and social ties.

By means of the integration of these steps we aim to enhance the inference related to the spatiotemporal dynamics of the users utilizing the social interactions between them, e.g., an user might go to a place because he has a friend there.

The developed model will be applied to the application scenario of mobile networks management for optimally and automatically implementing policies with the goal of meeting next generation telecommunication network performances. This will require the construction of a self-management system built on top of the behavioural model. The goal is to handle strategies related to data plane topology, flow forwarding, bandwidth requirements and propose personalized routing mechanism based on the extracted features. Moreover, text mining techniques will be used to identify users' interests and predict future requirements analyzing users' posts from online social networks.

III. Work in Progress

In this first part of the study, an analysis on human mobility has been carried out with the objective of estimating the future locations that users will visit. A predictor, which learns from user's mobility history and infers the next location, has been implemented employing a deep learning approach. In particular, a Recurrent Neural Network (RNN) fed by individual historical location data of the users, has been utilized in the estimation. The algorithm has been tested utilizing the Mobile Data Challenge (MDC) dataset [9], which involved 185 users in the data collection. Among all the available data present in the MDC dataset, only GPS traces have been utilized in the current state of the work. However, GPS data is noisy and need to be pre-processed to obtain accurate locations visited by mobile nodes. Our aim is to capture points, referred to as Points of Interest (PoIs), where a person is still or moving slowly. In order to define these locations we filter out points where users move faster than a threshold (set to 1.3m/s, according to the fact that human walking speed is about 1.1-1.4 m/s). After that, we utilize a Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to group together closely located points (less than 100m). At the end of these pre-processing phases the location history per each user is the temporal sequence of the PoIs. The predictor examines the history to predict the next location: the input space of the RNN is the history sequence and the target is the next PoI in the sequence. The training set is composed by the first 65% of the sequence, while the test set by the last 35%. A single hidden layer of 20 units has been employed in the RNN framework, achieving an average accuracy of 80.2%, which seems a prominent result according to the literature.

IV. Conclusion and Future Work

Human behaviour is a multifaceted problem, whose better understanding can be beneficial to several real-life aspects. We propose a stochastic analysis of data collected from social networks along with geospatial information to produce a model of human behaviour which aims to predict social dynamics and to infer users' mobility patterns and interests. A mobility predictor based on a RNN has been designed with the goal of estimating users' next visited locations. In the near future, this approach will be compared with other solutions and additional features will be included to increase the accuracy. Physical data will be coupled with social and contextual information to build the behavioural model. The proposed framework can open to improvement and optimization in several areas. As an application of the developed methodology we will consider the self-optimization of next generation telecommunication networks.

Acknowledgment

Supported by the SNSF via the SwissSenseSynergy project.

References

[1] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, 2010.

[2] M. Papandrea, K. K. Jahromi, M. Zignani, S. Gaito, S. Giordano, and G. P. Rossi, "On the properties of human mobility," *Computer Communications*, vol. 87, pp. 19–36, 2016.

[3] A. Förster, K. Garg, H. A. Nguyen, and S. Giordano, "On context awareness and social distance in human mobility traces," in *Mobile Opportunistic Networks*. ACM, 2012.

[4] J. L. Toole, C. Herrera-Yaqüe, C. M. Schneider, and M. C. González, "Coupling human mobility and social ties," *Journal of The Royal Society Interface*, vol. 12, no. 105, 2015.

[5] L. Backstrom, E. Sun, and C. Marlow, "Find me if you can: improving geographical prediction with social and spatial proximity," in *World Wide Web Conference*. ACM, 2010.

[6] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011.

[7] D. Hristova, M. J. Williams, M. Musolesi, P. Panzarasa, and C. Mascolo, "Measuring urban social diversity using interconnected geo-social networks," in *Proceedings of the 25th International Conference on World Wide Web*, 2016.

[8] A. Venerandi, G. Quattrone, L. Capra *et al.*, "Measuring urban deprivation from user generated content," in *Proceedings of the 18th Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015.

[9] J. K. Laurila, D. Gatica-Perez, I. Aad *et al.*, "The mobile data challenge: Big data for mobile computing research," in *Pervasive Computing*, no. EPFL-CONF-192489, 2012.