UNIVERSITÀ
DEGLI STUDI
DI PALERMO

Detection of Points of Interest in a Smart Campus

Article

Accepted version

A. De Paola, A. Giammanco, G. Lo Re, G, Anastasi

In Proceedings of I-CiTies 2019 - 5th Italian Conference on ICT for Smart Cities And Communities

Detection of Points of Interest in a Smart Campus

Alessandra De Paola, Andrea Giammanco, Giuseppe Lo Re,
University of Palermo
Palermo, Italy
{alessandra.depaola, andrea.giammanco, giuseppe.lore}@unipa.it

Giuseppe Anastasi
University of Pisa
Pisa, Italy
giuseppe.anastasi@unipi.it

***Abstract*—Understanding users' habits is a critical task in order to develop advanced services, such as personalized recommendation and virtual assistance. In this work, we propose a novel approach to detect Points of Interest visited by users of a campus, by using mobility traces collected through users' smartphones. Our method takes advantage of the intentional and recurrent nature of human movements to build up mobility profiles, and combines different machine learning methods to merge sensory information with the past users' behavior. The proposed approach has been validated on a synthetic dataset and the experimental results show its effectiveness.**

I. Extended Abstract

As the diffusion of personal devices increases, it is becoming easier to keep track of users' trajectories. In many fields, such as pervasive computing and social sciences, user profiling is a very valuable task, since it enables the creation of models of their activities, seen as sequences of movements. As proved in [1], human trajectories display a high degree of temporal and spatial regularity, and each user is characterized by a significant probability to return to a few highly frequented locations. Therefore, an accurate location identification makes location-aware applications more effective.

Context awareness is the key feature of *Smart Environments* [2], and in particular of a *Smart Campus* [3], [4], which is a digitally augmented campus where pervasive instrumented objects and spaces are made responsive to the state of the environment and its inhabitants. Location-aware computing aims at extracting information from raw trajectory data, in order to supply personalized services. In a Smart Campus, in addition to the ubiquity of users' smartphones, several other sensory devices, such as cameras, RFID readers, and bluetooth beacons, collect raw measurements, that can be exploited by an intelligent system in order to reason upon current context and supply advanced services to users. A location-based recommender system can provide information relevant to users' position, e.g. suggesting the nearest free library seat. Moreover, the capability of predicting users location can be exploited to provide recommendations related to the next place a user will visit, e.g., enabling the suggestion of free parking space near the next destination.

We propose a Smart Campus system which relies on a multi-tier architecture that allows to extract relevant information from raw data in order to provide context-aware services. Fig. 1 shows an overview of such architecture, with a detailed focus on the role played by the *Point of Interest (PoI) Detection* subsystem. A PoI is defined as a place where the user usually goes and stops for a while. The lowest layer of our architecture detects relevant events and monitors physical phenomena through a pervasive sensory infrastructure [5], [6]. Raw data gathered by such infrastructure, after a light preprocessing phase, are then exploited in order to perform the PoI detection. The *PoI Detection* subsystem analyzes raw data to detect a set of *Stop Points*, which can be defined as regions where the user remains for a certain amount of time, and merges the identified *Stop Points* with high-level context information in order to detect the actual visited PoIs. Such context information summarizes the relationship between users and the activities they perform in different PoIs, in terms of frequency of visit, average stop time, and sequence of visited PoIs. The detected PoIs are sent as input to the *Features Detection* module in order to extract a set of n-dimensional points which summarize the mobility behavior of each user, such as his leaning to be sedentary or exploratory during the workday. The *User Profiling* module then aims at outlining different classes of users according to previously extracted features. Besides inferred behavioral characteristics, user profiles also include information explicitly provided by the users themselves, such as the their role (e.g., "researcher" or "student"), which can be modeled through an opportune ontology [7]. Finally, knowledge about user profiles and their current and next position and activity can be exploited in order to provide users a set of *Context-Aware Services*.

The work described here mainly aims at inferring PoIs visited by users, using georeferenced position data. The regions where a user stops for a considerable amount of time can be automatically extracted by means of clustering [8], [9]. Nevertheless, mapping a user's position to a set of known PoIs is not a trivial task [10], [11], due to the intrinsic error in measurements and the presence of areas dense of meaningful places. We propose to face these issues, by adopting a combination of unsupervised and supervised machine learning methods, in order to be capable of dealing with uncertain data and of merging sensory data with knowledge about past user's behavior, which can be used to refine the estimation of their location. First, *Stop Points* are discovered through an enhanced version of the method described in [9], which integrates information stored in a database of known waypoints in order to merge consecutive *Stop Points* which match the same waypoint. Then, we propose to adopt a Dynamic Bayesian Network (DBN) [12] in order to include context awareness and the knowledge about past history into the PoI detection

DRAFT

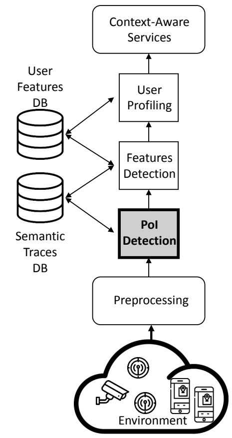

Fig. 1. Architecture of the proposed Smart Campus.

algorithm, and in order to deal with the unavoidable noise in sensory readings and in the discovered Stop Points.

Therefore, the extracted *Stop Points* represent the observable manifestation of the hidden user state, i.e., the real checked-in PoI. The DBN allows to model the evolution of the hidden state over time, and the probability dependency of the current state, from past state and from context features.

We propose to use a context information set which is not dependent on global features of PoIs, i.e. popularity or visibility on social media, but which heavily depends on the recurrent nature of human mobility [13], that can be often modeled through a weekly routine. According to this choice, our model takes simultaneously and separately into account three different aspects related to the time relation between the discovered *Stop Point* and the potential PoI: the arrival time at the *Stop Point*, the duration of the visit and the day of the week. Furthermore, we exploit the intentional nature of human mobility by considering the sequence of past visited places during the day as further contextual feature.

In order to test the effectiveness of the proposed PoI detection algorithm, we designed a simulation tool for building synthetic datasets of mobility traces, according to guidelines proposed in the literature [14]. The simulator represents each user as a set of commonly visited waypoints and a *behavior*, which is a probabilistic model for his specific mobility patterns. Such *behavior* is composed of two components: (i) a *transition scheme*, which is a Markov chain representing the transition probability from a waypoint to another, and (ii) a *pause scheme* which represents the probability distribution of stop time for each waypoint.

The first set of experiments carried out evaluates the performance of Stop Points extraction methods described in [8], [9] with our proposed version, under the ideal hypothesis of continous GPS signal. We perform the nearest neighbor assignment, considering the actually visited PoIs as those nearest to Stop Points' centroid, without further considerations. The overall accuracy of such methods can be evaluated by comparing the ground truth with the sequence of inferred PoIs through the Damerau-Levenshtein distance. We obtain an accuracy of 94%, thus increasing the accuracy of method [8] by 26% and of method [9] by 4%.

In the second experiment, we relax the constraint about the GPS signal continuity, which is now considered when a user moves inside buildings. We compare the performances of the blind nearest neighbor assignment method with our probabilistic approach with the probabilistic approach we propose, considering different subset of context information. The training and validation of our DBN model have been performed by a 6-fold cross validation on the whole data set. Results show that optimal performances are achieved by considering user's arrival time and stay duration, obtaining an accuracy of 90%, thus increasing the accuracy of nearest neighbor assignment by 7%.

As future work, we will address other modules of the proposed context-aware system, also by including the possibility of taking into account users' activity detected from pervasive sensors as further context feature [15], [16].

Moreover, we plan to engage users in labeling PoIs' classes, through a participatory sensing approach that relies on techniques of reputation management [17] to discard information provided by malicious users who intentionally send incorrect data to create a disservice.

Acknowledge

This research is partially funded by the Project VASARI of Italian MIUR (PNR 2015-2020, DD MIUR n. 2511).

References

[1] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *nature*, vol. 453, no. 7196, p. 779, 2008.

[2] A. De Paola, P. Ferraro, S. Gaglio, G. Lo Re, M. Morana, M. Ortolani, and D. Peri, "A context-aware system for ambient assisted living," in *Proc. of Int. Conf. on Ubiquitous Computing and Ambient Intelligence*, 2017, pp. 426–438.

[3] F. Concone, P. Ferraro, and G. L. Re, "Towards a smart campus through participatory sensing," in *Proc. of the 2018 IEEE Int. Conf. on Smart Computing (SMARTCOMP)*, 2018, pp. 393–398.

[4] P. Agate, F. Concone, and P. Ferraro, "Wip: Smart services for an augmented campus," in *Proc. of the 2018 IEEE Int. Conf. on Smart Computing (SMARTCOMP)*, 2018, pp. 276–278.

[5] A. De Paola, P. Ferraro, S. Gaglio, G. Lo Re, and S. K. Das, "An adaptive bayesian system for context-aware data fusion in smart environments," *IEEE Trans. on Mobile Computing*, vol. 16, no. 6, pp. 1502–1515, 2017.

[6] A. De Paola, P. Ferraro, S. Gaglio, and G. Lo Re, "Context-awareness for multi-sensor data fusion in smart environments," in *Proc. of the Conf. of the Italian Association for Artificial Intelligence*, 2016, pp. 377–391.

[7] S. Gaglio and G. Lo Re, *Advances onto the Internet of Things*. Springer, 2014, vol. 260.

[8] R. Montoliu, J. Blom, and D. Gatica-Perez, "Discovering places of interest in everyday life from smartphone data," *Multimedia Tools and Applications*, vol. 62, no. 1, pp. 179–207, 2013.

[9] Z. Fu, Z. Tian, Y. Xu, and C. Qiao, "A Two-Step Clustering Approach to Extract Locations from Individual GPS Trajectory Data," *ISPRS International Journal of Geo-Information*, vol. 5, p. 166, 2016.

[10] K. Nishida, H. Toda, T. Kurashima, and Y. Suhara, "Probabilistic identification of visited point-of-interest for personalized automatic check-in," in *Proc. of the 2014 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing (UbiComp 2014)*, 2014, pp. 631–642.

[11] B. Shaw, J. Shea, S. Sinha, and A. Hogue, "Learning to rank for spatiotemporal search," in *Proc. of the 6th ACM Int. Conf. on Web Search and Data Mining (WSDM 2013)*, 2013, pp. 717–726.

DRAFT

[12] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.
[13] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, “Limits of predictability in human mobility,” *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
[14] F. Giurlanda, P. Perazzo, and G. Dini, “Humsim: A privacy-oriented human mobility simulator,” in *International Conference on Sensor Systems and Software*. Springer, 2014, pp. 61–70.
[15] F. Concone, G. Lo Re, and M. Morana, “A fog-based application for human activity recognition using personal smart devices,” *ACM Trans. on Internet Technology (TOIT)*, vol. 19, no. 2, p. 20, 2019.
[16] F. Concone, S. Gaglio, G. Lo Re, and M. Morana, “Smartphone data analysis for human activity recognition,” in *Proc. of the Conference of the Italian Association for Artificial Intelligence*, 2017, pp. 58–71.
[17] V. Agate, A. De Paola, G. Lo Re, and M. Morana, “A simulation framework for evaluating distributed reputation management systems,” in *Proc. of the 13th International Conference on Distributed Computing and Artificial Intelligence*. Springer, 2016, pp. 247–254.

DRAFT