A Spatiotemporal Approach for Social Situation Recognition

Christian Meurisch(✉), Tahir Hussain, Artur Gogel, Benedikt Schmidt, Immanuel Schweizer, and Max Mühlhäuser

Telecooperation Lab, Technische Universität Darmstadt, Darmstadt, Germany
{christian.meurisch,tahir.hussain,artur.gogel,
benedikt.schmidt,schweizer,max}@tk.informatik.tu-darmstadt.de

Abstract. The development of virtual personal assistants requires situation awareness. For this purpose, lightweight approaches for the processing of sensor data to derive situation information from available sensor data (e.g., mobile phone data) are required.

In this paper, we propose a spatiotemporal approach to derive situational information about social interactions only based on location and time, using data collected with off-the-shelf smartphones. We examine the approach, using location traces of 163 users collected over four weeks. The proposed spatiotemporal approach shows an average social situation recognition result of $45.8 \pm 23.2\,\%$ F_1-measure across the data set using Random Forest classifiers.

Keywords: Social interaction · Personal tracking · Mobility pattern · Social computing · Situation recognition · Location sensing · Smartphone

1 Introduction

Virtual personal assistants to support users in their daily lives have become more and more popular in recent years. Due to the growing use of mobile devices assistant systems are able to seamlessly track and give advices through mobile applications at any time. Examples are commercial personal assistants like Google Now or Apple's Siri which offer automatic event reminders considering related information like traffic situations. Similarly, fitness trackers offer coaching considering users' actual performance [16], can detect and even prevent health problems [4]. In all cases knowledge about the user's current situation is of utmost importance.

A major challenging problem in situation recognition is the information base. In general the fusion of already available information sources (e.g., mobile sensor data, web data) [5,20] is preferred over deploying additional static hardware (e.g., smart home sensors) [1]. After data gathering, the fusion and processing of sensor data to derive information types, granularity and quality suitable for situation follows. We propose a lightweight approach to derive social interactions based only on location and time with no additional instrumentation of the user or the environment.

S. Sigg et al. (Eds.): MobiCASE 2015, LNICST 162, pp. 309–316, 2015.
DOI: 10.1007/978-3-319-29003-4_18

In this paper, we examine over 24 million location traces of 163 students over four weeks to automatically infer the user's *social situation* represented by place, time and social presence. For that, we reduce the complex situation recognition problem to the detection of social interactions from mobility patterns. In our approach, a *social interaction* is defined as meeting of a group consisting of at least two persons who are co-located over a specific amount of time (e.g., five minutes). By doing this, our data model consists of three types of information: (1) *social*, (2) *spatial*, and (3) *temporal information*. We assume that each one of the three information types can be derived by using the other two. This paper focuses on inferring social interactions (class label) from spatiotemporal data (features), i.e., if we know the user's current whereabouts and the corresponding times, we are able to infer information about the social situation which means we know the exact persons a social interaction takes place with. While location and time are standard information of modern off-the-shelf smartphones, our approach is highly suitable for daily use. Training personalized classifiers for each user we get an overall classification result from $45.8 \pm 23.2\,\%$ F_1-measure across our user base.

In summary, the contributions of this paper are twofold:

Self-tracking Dataset. Using personal mobile devices we collected a large dataset with 24 million location values of 163 students over four weeks using our multi-device user tracking suite [17,18].

Social Situation Recognizer. Detecting social interactions only from location traces, we extract features and train personalized Random Forest classifiers for each user [3]; averaging the results over all users we get an overall classification result of $45.8 \pm 23.2\,\%$ F_1-measure across our users. Thus, we are able to infer the current user's social situation from spatiotemporal data stream.

The remainder of this paper is organized as follows. In Sect. 2, we provide an overview of related work in recognition users' activities and contexts. In Sect. 3, we present our approach that utilizes detection of social interactions to reveal user's social situation. In Sect. 4, the data collection process with the resulting dataset and feature extraction process are described, before we report the results. The paper closes with conclusion and future works.

2 Related Work

Most mobile phones feature a large variety of sensors [13], providing the perfect platform for activity, context and situation recognition [22] without the need of deploying additional hardware like in smart homes [1]. While simple activities (e.g., standing, walking) can be detected with a high accuracy of above 90 % relying on accelerometer data only [12], the recognition of complex activities (e.g., watching tv, playing volleyball) or detecting the entire situation including social presence, only with highly available devices is still an open challenge [15]. Most approaches still need external sensors or custom wearables [9].

Our idea is based on only utilizing location traces of users to recognize social interactions (i.e., temporal co-located users) and characterize the user's current situation. Other state of the art approaches already utilize the social context inferred from mobility pattern [6] or online social networks [5] to improve their results in human activity recognition [14]. In this paper, we focus on one data source deployed in all modern mobile phones (i.e., locations) to underpin that deep insights into activity and situation can be derived only by location traces. Various works already show recognition approaches based on location traces, e.g., place detection [11,23], social relationship inferences [21] or even the recognition of mental disorders like depressions [4]. It is important to note that human trajectories show a high degree of temporal and spatial regularity [8]. Song et al. even find a 93 % potential predictability in user mobility across their user base [19], which makes human mobility an attractive data source for context and situation recognition. In the next section, we describe our spatiotemporal approach in detail.

3 Our Approach

In this section we describe our approach for deriving a current social situation (the persons interacted with) from space and time (cf. Fig. 1). For that, we only need the location sensor of the smartphone to get location (coordinates) or location traces of a user, including timestamps. In the following we explain our approach of inferring social interactions from human mobility step by step. *First*, we describe the detection of places (high-level spatial information) and place visits (high-level spatial-temporal information) inferred from location traces (low-level information) for each user. *Second*, we cluster temporal overlapping place visits of users to social interactions (high-level spatial-temporal social context). *Finally*, we represent a situation by these three information: (1) *social*, (2) *spatial*, and (3) *temporal data*.

3.1 Places and Place Visits

We define a *place* as stationary geographical location where a user stays for an amount of time. We extract a user's places from location traces by using the place detection algorithm proposed by [11] with $d_p = 25\,\mathrm{m}$ and $t_p = 15\,\mathrm{min}$ as distance and time threshold parameters. Within the detected places, we identify places of specific relevance throughout the daily live. Inspired by [23], we define and determine the following four meaningful place types: *home*, *work*, *university*, and *other place* (high-level spatial information). To add temporal information to that static places, we check when and how long users visit their specific places and define this as *place visit* (high-level spatial-temporal information).

3.2 Social Interactions

We define a *social interaction* as a meeting of at least two persons who are temporal co-located over a specific amount of time (*here:* at least t_s =5 min).

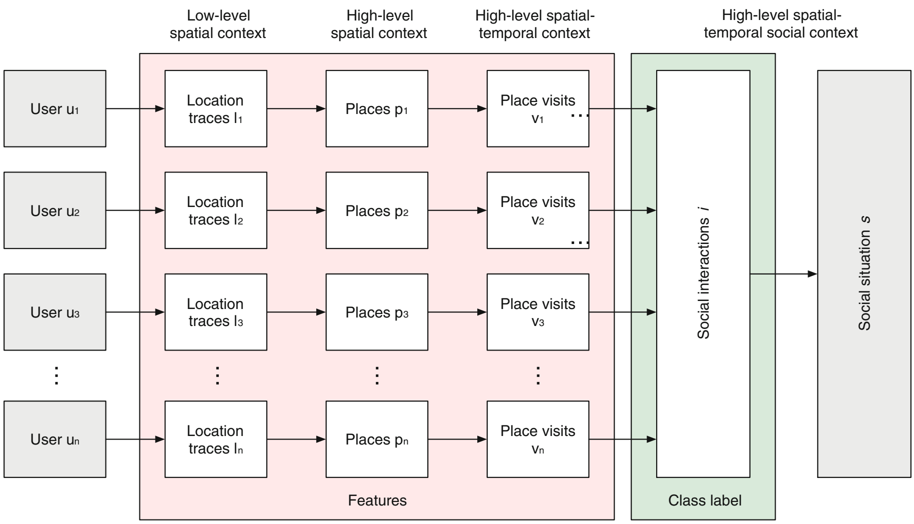

Fig. 1. Our approach utilizes location traces of users to detect higher-level information like places, place visits (features) and, finally, social interactions (class label) to characterize a situation.

Inspired by [5,6,21], we infer social interactions from co-location of users. More precisely, we consider spatiotemporal overlapping place visits of different users in a sliding window with size of t_s to detect social interactions (high-level spatial-temporal social context). For that, we use the clustering algorithm DBSCAN [7] with parameters $minPts = 2$ and eps =20 m within that sliding window to cluster users to social groups.

3.3 Social Situation Recognition

A *social situation* in our approach is represented by three contexts: (1) *social*, (2) *spatial*, and (3) *temporal* context. Given a use case, we are able to sense two of these contexts and infer the third context to characterize the situation. In this paper, we focus on the recognition of the social context from spatiotemporal data. For example, if we are able to sense location values including a timestamp, our approach can infer social interactions (cf. Fig. 1), i.e., the output is the exact user group with unique persons. Utilizing well-researched *next place detection* algorithms [2], our approach could furthermore be utilized to predict the exact persons a user will meet the next or in near future.

4 Proof of Concept

To prove our approach we first conducted a user study to get real-world location data, described in the next section. Based on that dataset we extract appropriate spatiotemporal and social features (class label) to train a personalized classifier for each user to recognize user's social situations. Finally, we report and discuss the classification results.

4.1 Dataset

We conducted a self-tracking user study to collect location data from 163 students of Technische Universität Darmstadt over four weeks using our multi-device user tracking suite [17,18]. In total, we gathered over 24 million raw location values within four weeks, i.e., about 148 ± 359 thousand location values per user. The high scatter can be reasoned by dynamic sampling rates for location sensor depending on the strength of user's movement, i.e., we reduce the sampling rate if the smartphone is still, while we increase the sampling rate if the smartphone is moving, especially in vehicles.

Table 1. Obtained higher-level information from collected locations

Step	High-level information	Instances per user	Total
-	Raw locations l	$147,630.7 \pm 358,764.9$	$24,063,641$
1	Places p	14.2 ± 10.8	$2,312$
2	Place visits v	102.0 ± 59.7	$16,629$
3	Social interactions i	182.7 ± 193.7	$29,787$

Table 2. Definition of spatiotemporal features (f_p,f_w,f_d,f_t) and class label (f_s)

ID	Feature	Value range
f_p	Place type	{home, work, university, other}
f_w	Weekend	{false, true}
f_d	Day of week	{Mon, Tue, Wed, Thu, Fri, Sat, Sun}
f_t	Time of day	{morning, afternoon, evening, night}
f_s	Social interaction	$\{u_k, u_j, ..\} \subseteq U$ (users)

4.2 Feature Extraction

For the dataset, three different kinds of features were extracted: (1) *social*, (2) *spatial*, and (3) *temporal features*. For that, we proceed as described in the previous section: inferring places from raw location values (*step 1*), detecting place visits (*step 2*), and clustering social interactions (*step 3*). Table 1 shows the resulting count of instances per user and the total count of instances for each processing step. Based on this high-level information the features are extracted. Table 2 lists the resulting features and their value ranges. Feature f_p represents the place context with four possible semantic places for each user: *home*, *work*, *university* and *other places*. Finally, we have three time-based features: the binary feature f_w (weekend) with *false* for weekday or *true* for weekend as values; feature f_d (day of week) with the seven days of week as values (i.e., Monday, Tuesday,..), and feature f_t (time of day) with value range of *morning* ($6am - 12pm$), *afternoon*($12pm - 17pm$), *evening* ($17pm - 22pm$), and *night* ($22pm - 6am$). As categorical class label, we use the feature f_s representing the social context. Its value range is an arbitrary subset of all users within the system, i.e., the smallest subset contains only the user himself and the largest subset contains all users. In total, we extracted 59,412 instances, i.e., 364.5 ± 237.0 instances per user.

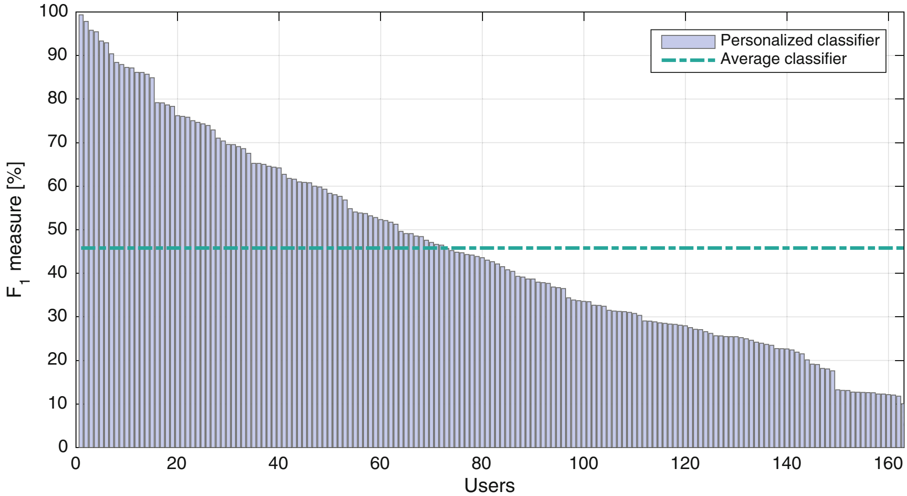

Fig. 2. Classification results (F_1-measure) of personalized classifiers for each users (*indigo*) and the average F_1-measure over all users (*green line*) (Color figure online)

4.3 Results

With the above extracted spatiotemporal features (f_p,f_w,f_d,f_t) and f_s as class to predict we train and evaluate personalized classifiers for each user. For that, we programmatically tested 26 various classification algorithms with different configurations provided by WEKA, a data mining software [10]. Avoiding overfitting, the best evaluated classifier was a *Random Forest* with 100 trees [3]. In Fig. 2, we report the results for each personalized classifier and the average F_1-measure of $45.8 \pm 23.2\,\%$ over our user base using the above classification algorithm. We see that social context (i.e., exact determination of present users) is highly predictable for few users, i.e., F_1-measure ranging between 70.0 % and

99.3 % for about 18 % of users. For over 38 % of users our approach is able to correctly detect the exact social interactions in every second situation. For the rest of users the social context prediction out of spatiotemporal data is challenging. In future work, we plan to assign each probably presented unique person a probability of attendance, i.e., the algorithm will consider subgroups, to further improve our results.

5 Conclusion

In this paper, we proposed a recognition approach to detect user's social situation only utilizing his location traces. By analyzing these location traces we recognize social interactions (i.e., co-located persons over a specific time) to derive the social situation. Our evaluation built on a four-week self-tracking study with over 24 million location values of 163 students. We showed an average recognition result of $45.8 \pm 23.2\,\%$ F_1-measure across our user base using personalized Random Forest classifiers for each user. The result confirms the initial assumption that if the same group of users meets each other, the situation with respect to time and place is often the same or similar and, thus, predictable. Therefore, the presented approach can predict the persons of prospective meetings. In future work, we will also consider the relationship between social interactions, e.g., friends, classmates or colleagues, and investigate the impact of user routines versus accidental meeting. Moreover, we will build a real-time assistive system based on our situation recognition approach to support students in their daily lives.

Acknowledgments. This work has been funded by the LOEWE initiative (Hessen, Germany) within the NICER project.

References

1. Arcelus, A., Jones, M.H., Goubran, R., Knoefel, F.: Integration of smart home technologies in a health monitoring system for the elderly. In: 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW 2007), vol. 2, pp. 820–825. IEEE (2007)
2. Baumann, P., Kleiminger, W., Santini, S.: The influence of temporal and spatial features on the performance of next-place prediction algorithms. In: 15th International Conference on Ubiquitous Computing (UbiComp 2013), pp. 449–458. ACM (2013)
3. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)
4. Canzian, L., Musolesi, M.: Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: 17th International Conference on Ubiquitous Computing (UbiComp 2015), pp. 1293–1304. ACM (2015)
5. Cranshaw, J., Toch, E., Hong, J., Kittur, A., Sadeh, N.: Bridging the gap between physical location and online social networks. In: 12th International Conference on Ubiquitous Computing (UbiComp 2010), pp. 119–128. ACM (2010)

6. Eagle, N., Pentland, A.: Reality mining: sensing complex social systems. Pers. Ubiquit. Comput. **10**(4), 255–268 (2006)
7. Ester, M., Kriegel, H.-P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: 2th International Conference on Knowledge Discovery and Data Mining (KDD 1996), pp. 226–231. ACM (1996)
8. Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.-L.: Understanding individual human mobility patterns. Nature **453**(7196), 779–782 (2008)
9. Gordon, D., Hanne, J.-H., Berchtold, M., Shirehjini, A.A.N., Beigl, M.: Towards collaborative group activity recognition using mobile devices. Mob. Netw. Appl. **18**(3), 326–340 (2013)
10. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
11. Kang, J.H., Welbourne, W., Stewart, B., Borriello, G.: Extracting places from traces of locations. Mob. Comput. Commun. Rev. **9**(3), 58–68 (2005)
12. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. Explor. Newsl. **12**(2), 74–82 (2011)
13. Lane, N.D., Miluzzo, E., Hong, L., Peebles, D., Choudhury, T., Campbell, A.T.: A survey of mobile phone sensing. Commun. Mag. IEEE **48**(9), 140–150 (2010)
14. Lane, N.D., Pengyu, L., Zhou, L., Zhao, F.: Connecting personal-scale sensing and networked community behavior to infer human activities. In 16th International Conference on Ubiquitous Computing (UbiComp 2014), pp. 595–606. ACM (2014)
15. Lara, O.D., Labrador, M.A.: A survey on human activity recognition using wearable sensors. Commun. Surv. Tutorials **15**(3), 1192–1209 (2013)
16. Schmidt, B., Benchea, S., Eichin, R., Meurisch, C.: Fitness tracker or digital personal coach: how to personalize training. In: 17th International Conference on Ubiquitous Computing (UbiComp 2015): Adjunct Publication. ACM (2015)
17. Schweizer, I., Bärtl, R., Schmidt, B., Kaup, F., Mühlhäuser, M.: Kraken.me mobile: the energy footprint of mobile tracking. In: 6th International Conference on Mobile Computing, Applications and Services (MobiCASE 2014), pp. 82–89. IEEE (2014)
18. Schweizer, I., Schmidt, B.: Kraken.me: multi-device user tracking suite. In: 16th International Conference on Ubiquitous Computing (UbiComp 2014): Adjunct Publication, pp. 853–862. ACM (2014)
19. Song, C., Zehui, Q., Blumm, N., Barabási, A.-L.: Limits of predictability in human mobility. Science **327**(5968), 1018–1021 (2010)
20. Takata, K., Ma, J., Apduhan, B.O., Huang, R., Shiratori, N.: Lifelog image analysis based on activity situation models using contexts from wearable multi sensors. In: 2nd International Conference on Multimedia and Ubiquitous Engineering (MUE 2008), pp. 160–163. IEEE (2008)
21. Wang, D., Pedreschi, D., Song, C., Giannotti, F., Barabasi, A.-L.: Human mobility, social ties, and link prediction. In 17th International Conference on Knowledge Discovery and Data Mining (KDD 2011), pp. 1100–1108. ACM (2011)
22. Yau, S.S., Liu, J.: Hierarchical situation modeling and reasoning for pervasive computing. In: 4th Workshop on Software Technologies for Future Embedded and Ubiquitous Systems, and 2nd International Workshop on Collaborative Computing, Integration, and Assurance (SEUS-WCCIA 2006), p. 6. IEEE (2006)
23. Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L.: Discovering personally meaningful places: an interactive clustering approach. ACM Trans. Inf. Syst. (TOIS) **25**(3), 12 (2007)