Predicton With Smart Phones

Gowtham Mamidisetti, B.Venkatesh

Abstract--This paper predicts two aspects of human behavior using smart phones as sensing devices. This paper introduces a method for predicting where users will go and which application they will use next by exploiting the rich contextual information from smart phone sensors. Our first goal is to understand which smart phone sensor data types are important for the two prediction tasks. Secondly, we aim at extracting user independent behavioral patterns and study how user independent behavior models can improve the predictive performance.

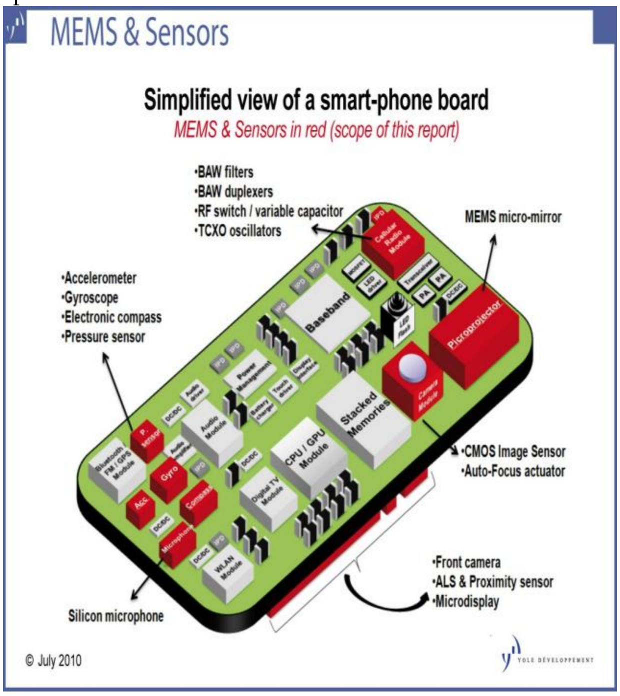
Keywords: smart phone data, human behavior, mobility prediction, app usage prediction.

I. INTRODUCTION

The ability to predict what next mobile users want to do has many applications such as improving user interfaces or providing relevant recommendations. For example, if the system knows that the user will go for lunch, then it can recommend a list of restaurants together with useful information like today's menu, availabilities, traffic conditions, etc. Mobile phones are convenient options for tracking and mining user behavior in daily life as they are usually placed in close proximity to the users. Smart-phones contain many sensors that can record activities, including location, application usage, and calling behavior. This information can be considered as both input and output of a prediction method, in which the future values of some variables are predicted based on the current context .

In this paper, we consider the task of predicting human behavior based on multiple smart phone sensors using statistical methods. we predict the next location of a user and which application he/she will use based on the current context consisting of location, time, app usage, Bluetooth proximity, and communication logs.

This approach allows modeling the interplay between the predicted variables to study relationships between the place where a user stays and the possibility that he would make a phone call, use the cameras and so on. Furthermore, other sources of information, such as the list of nearby Bluetooth devices or system information, are also exploited in order to enrich the user context and improve the predictive models. First, we present a general framework for predicting jointly various dimensions of human behavior using multimodal data. Second, we study the impact of each data type to the predictive performance of the two tasks. Third, we investigate the use of generic behavior patterns for improving prediction performance of personalized models.



II.METHODOLOGY

we first describe the dataset used in our study. Next, we describe the task that we address along with various data representations, which will be used throughout this paper. Finally, we describe various statistical models that can be integrated in our prediction based on the information we get from smart phone sensors.

1. The dataset

We use data collected with phones and a continuous sensing software capable of periodically saving sensor data from volunteers. This enabled us to collect long-term data from several modalities such as GPS, Bluetooth (BT), WiFi access points, accelerometer, and applications and phone usage logs. The software contained a state-machine approach similar to define an adaptive sensing procedure. Operating states (like "indoors", "outdoors", "travel", etc.) were determined heuristically to maximize battery life by sampling from the sensors

Manuscript Received November, 2013.

Gowtham.Mamidisetti is an assistant Professor in Information Technology at Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India.

B.Venkatesh is an assistant Professor in Information Technology at Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India.

at different rates depending on the operating states.
Location data preprocessing from the raw location coordinates extracted by the sensing software through GPS and known WiFi-access points, se- mantic labels were assigned to the most common places using a two-step procedure . First, we discovered places that a user visited repeatedly or for a long duration of time, and then self-reported annotations of these places from the users themselves were obtained. Places are defined as small circular regions in which users stay for at least 10 minutes. The radius of place is set to 100 meters to deal with the existence of noisy data at some location.

In the second step, we annotate these stay regions into a set of semantically meaningful labels .we annotated eight regions on a map through a web interface .major regions corresponded to the most frequently visited places by the user.
In our method, we also use application usage, Bluetooth proximity, communication statistics, and operating state of the sensing software. The user context is then represented by several variables of multiple modalities.
Since people often used a limited number of applications .we selected only the most frequently used applications for building the user context. The application logs consist of usage events of all applications, included system applications and preinstalled apps like Camera or Clock, and user downloaded applications like Gmail.

2. Data representation

We use data streams to encode the human behavior over time and use statistical methods to learn a predictive human behavior model. We first group together location labels, applications usage, and other contextual information into 10 minute time slots. This enables us to view smart phone data in the form of a set of temporally arranged data streams, with each stream corresponding to one modality such as GPS . This is conceptualized into Several of these streams can now be grouped into categories. In our method, of the many possibilities, we concern ourselves with six categories, that are explained below.

1.Location: Consisting of 10 binary data streams where 10 is the number of semantic location labels. Each stream corresponds to a significant labeled place that a user has been found to visit .

2. Apps: Consisting of the set of app streams, where each stream
corresponds to the count of app-view-events of an app that is used least once a week by the user . Since each user used a different set of apps, the number of app streams varies for different users.

3. Time: We use time-of-day and day-of-week to represent the temporal context.

4.Call log :Consisting of data 5 binary streams representingthe number of communication events for each time slot. The 5 data streams corresponds to 5 categories of communication event: SMS received, SMS sent, outgoing call, incoming call, and missed call.

5. Bluetooth proximity: From BT scans, we extract the number of nearby BT devices and the presence of frequently observed BT devices. .

6 Internal state: Internal operating states of the recording software are also used as context. These states are inferred with several heuristic rules on accelerometer data, WiFi readings, GPS, and system information.

These categories together describe the overall usage of smart phone by a user.

III. MOBILITY PREDICTION:

1. Generic human mobility patterns
   We use correlation analysis to characterize the dependency between contextual variables and the next location. A strong correlation indicates a high predictability degree of the output using a linear model on the considered contextual variable.
2. Personalized mobility prediction
   Accuracy. While generic human mobility can be captured, people have their "own" routines in real life. Therefore, personalized predictive models are expected to be more accurate than generic model. since the most frequent location stream changes depending on the user.

Finally, combining the generic model with the personalized model increases the overall accuracy , which shows the usefulness of generic model on the prediction

IV. APP USAGE PREDICTION

1. Generic app usage patterns

One common observation is that the use of an app at a given time is highly correlated with the use of that app in the next 10 minutes. For example, in the case of Camera , if the user is using camera then there is a high probability that the user continues to take more photos/videos in the next ten minutes. Sim- ilar to the case of mobility correlation analysis, most of correlation values are small but statically significant . We can find many meaningful results such as Text message is likely to be opened in the next 10 minutes (to send or to receive messages) if the user recently sent/received SMS. Another interesting example is the case of Profiles app which is used for changing ringing types of the phone.

2. Personalized app usage prediction

we study the effect of number of apps on the performance and on the improvement of context- aware model over the baseline (MostFrequent). the recall values and the improvement of the Personalized+Generic method over the baseline, and the number of candidate apps for each user. To track the tendency of performance with respect to the number of apps, we used a second order polynomial to fit the data.

V. CONCLUSION

In this paper, we have utilized large-scale, long-term longitudinal smart- phone data to predicti two aspects of user behavior, i.e., location and app usage. We integrate the rich multimodal information made available through the smartphone sensors to predict location and application usage of the user, based on past and current data.

In the case of predicting location, we found that Bluetooth proximity is an important contextual cue along with Location and Time. This finding confirms again the dependency between human mobility and social interac- tions. In the second prediction task, the potential of our method to infer application usage in the future. To our knowledge, little work has been done to predict applications usage based on past activities

REFERENCES

1. G. Adomavicius, A. Tuzhilin, Context-aware recommender systems, Recommender Systems Handbook (2011) 217–253.
2. T. Liu, P. Bahl, I. Chlamtac, Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks, Selected Areas in Com- munications 16 (6) (1998) 922–936.
3. J. Krumm, E. Horvitz, Predestination: Inferring destinations from par- tial trajectories, in: Proc. UbiComp, 2006, pp. 243–260.
4. N. Eagle, A. (Sandy) Pentland, Reality mining: sensing complex social systems, Personal and Ubiquitous Computing 10 (4) (2006) 255–268.
5. A. Oulasvirta, R. Petit, M. Raento, S. Tiitta, Interpreting and acting on mobile awareness cues, Human–Computer Interaction 22 (1-2) (2007).
6. S. N. Patel, J. A. Kientz, G. R. Hayes, S. Bhat, G. D. Abowd, Farther than you may think: An empirical investigation of the proximity of users to their mobile phones, Ubiquitous Computing 4206 (2006) 123–140.
7. K. Farrahi, D. Gatica-Perez, Discovering routines from large-scale hu- man locations using probabilistic topic models, Transactions on Intelli- gent Systems and Technology 2 (1) (2011) 3.
8. A. Peddemors, H. Eertink, I. Niemegeers, Predicting mobility events on personal devices, Pervasive and Mobile Computing 6 (2010) 401–423.

AUTHOR PROFILE

Gowtham.Mamidisetti is an assistant Professor in Information Technology at Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India.

B.Venkatesh is an assistant Professor in Information Technology at Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India.