Analyzing the Influence of Phone Context Data on the Performance of Human Mobility Predictors

Paul Baumann, Silvia Santini
Embedded Systems Lab, TU Dresden, Germany
{paul.baumann, silvia.santini}@tu-dresden.de

Christian Koehler, Anind Dey
Carnegie Mellon University, USA
{ckoehler, anind}@cs.cmu.edu

1. INTRODUCTION

Understanding and predicting human mobility has been already for a long time in the focus and interest of researchers and practitioners [4, 2, 5]. Over the years, the potential sources of information about human mobility, e.g., temporal data, calendar information, or social ties, have been grown dramatically.

In the context of this extended abstract, we focus on the phone context data, e.g., number of recently used applications or time since the last received or made phone call. By considering this new source of data, we analyze its influence on the performance of 4 state-of-the-art and 3 baseline predictors for 3 human mobility prediction tasks. We capture the results of the potential influence by considering 3 well-known performance metrics. Our results highlight that our use of the phone context data *does not* lead to signification performance improvements. Statistical information from the data set shows that users used 7 applications, and made/received 5 phone calls, on average per day. This highlights the fact that the phones have been used intensively by their owners.

Our contributions are two-fold: (1) we derive an extended list of 28 phone context features, and (2) we analyze their potential influence on all combinations of the considered 3 prediction tasks, 7 predictors, and 3 performance metrics.

2. METHODOLOGY AND BACKGROUND

In order to investigate the influence of phone context data on the performance of several predictors for human mobility, we analyze a rich data set – Nokia Lausanne Data Collection Campaign (LDCC) [3] – that contains information from 141 users collected over 18 months. We first derive a list of meaningful phone context features based on the available information in the data set. After that, we run feature selection to reduce the amount of derived features for all combinations of predictors, prediction tasks, and metrics. We evaluate the performance of the resulting combinations of a predictor and a set of features on well-known metrics – accuracy, F1 score, and Matthews Correlation Coefficient (MCC).

2.1 Features

The human mobility predictors that are used throughout this work need adequate input data to be able to compute the required prediction. To this end, we derive a list of 51 features that contains 28 phone context features – indicated as set $\mathcal{F}_{pc}$. Table 1 shows them along with their corresponding description.

2.2 Prediction Tasks

In the context of this work, we consider 3 prediction tasks. The *Next-place (NP)* prediction task only considers the next place visited and treats as irrelevant when the user moves to the next location and how long she stays in each place. Timing information can easily be included in a prediction task by considering equally spaced time slots of length s and computing a new next place prediction for each time slot. We refer to this task as the *Next-slot place (NSP)* prediction task. Finally, the *Next-slot transition (NST)* prediction task consists of estimating, at time slot k, whether or not there will be a *transition* at time slot $k+1$. A transition occurs when the user moves between two places.

[1] We consider calls and messages for both incoming and outgoing directions.

[2] We consider date of creation, status, title, location, type, and confidence class.

Table 1: List of phone-context features considered in this study.

Label	Description
c_time_call[1]	Time since last call/sms made/received
c_calllog_type	Last calllog type
c_calllog_direction	Last calllog direction
c_sms_status	Last sms status
c_last_call_duration	Last call duration
c_last_cal[2]	Information about last calendar entry
c_next_cal[2]	Information about next calendar entry
c_time_last_app	Time since last application used
c_last_app	Last used application
c_phone_charging	Current phone charging status
c_last_charge	Time since last charge
c_battery	Current phone battery status
c_ring	Current ring profile
c_profile	Current user profile
c_last_action	Time since last phone interaction

2.3 Predictors

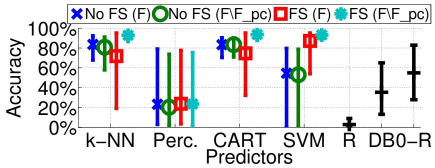
In this study, we consider 7 predictors that rely on different basic techniques, have different weaknesses and strengths, and require different amounts of computational and memory resources. The 7 predictors include 4 well-known predictors – Support Vector Machine (SVM) [6, 1], k-Nearest Neighbor (k-NN) [6, 1], Classification and Regression Trees (CART) [6], and Perceptron [1] – as well as 3 baseline predictors – Random predictor (R), Distribution-based predictor (DB), and 0-R predictor (0-R).

3. INFLUENCE OF PHONE CONTEXT DATA ON PREDICTORS' PERFORMANCE

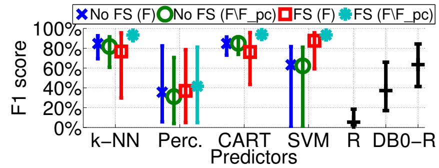
For our analysis part, we adopt the Sequential Forward Floating Selection (SFFS) algorithm to identify the best performing features for all combinations of prediction tasks, metrics, and predictors. In all these cases, we differentiate between using the entire feature set $\mathcal{F}$ and using only a reduced feature set by excluding phone context features ($\mathcal{F} \setminus \mathcal{F}_{pc}$). We next present our findings.

3.1 Performance Results

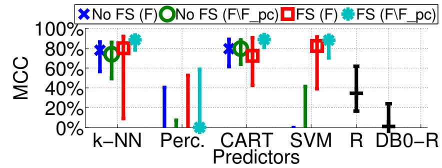
The feature selection process returns an optimal feature subset $\mathcal{F}'$ for each combination of user, predictor, prediction task, and metric. Figure 1 shows the performance in terms of accuracy, F1



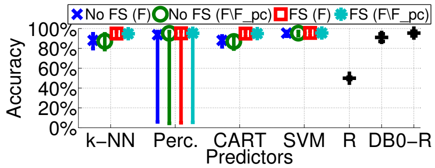
(a) Next-slot place – Accuracy



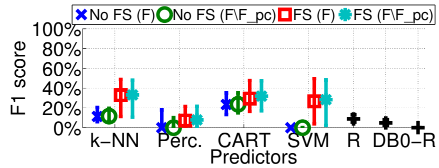
(b) Next-slot place – F1 score



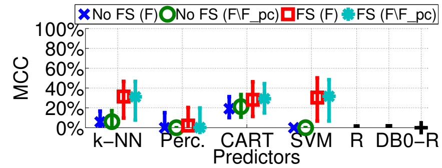
(c) Next-slot place – MCC



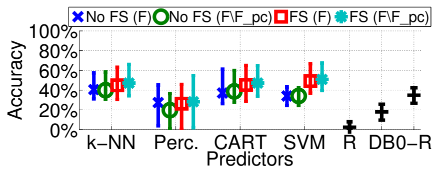
(d) Next-slot transition – Accuracy



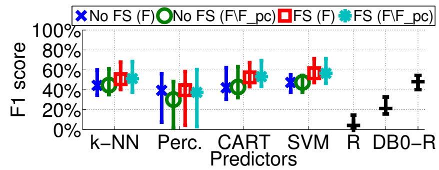
(e) Next-slot transition – F1 score



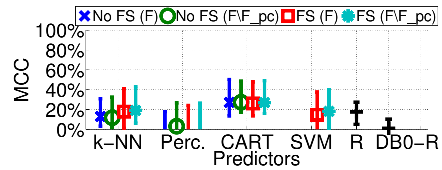
(f) Next-place transition – MCC



(g) Next-place – Accuracy



(h) Next-place – F1 score



(i) Next-place – MCC

Figure 1: Performance results for all 3 considered prediction tasks, 3 performance metrics, and 7 predictors.

score, and MCC achieved by the predictors with or without the feature selection ("FS" vs "No FS"). The markers indicate median values and the whiskers indicate the 5th and 95th percentiles.

Figure 1a, Figure 1b, and Figure 1c show the median performance for the NSP prediction task. The plots for the metrics accuracy and F1 score reveal significant performance differences for the extracted feature sets with phone context features ($\mathcal{F}_{pc}$) and those without. It holds for the predictors k-NN and CART. In these cases, the performance after applying feature selection is higher for the feature set $\mathcal{F} \backslash \mathcal{F}_{pc}$ than for $\mathcal{F}$. It is also the case for MCC, but with a lower performance difference in the case k-NN is used. Furthermore, SVM shows similar performance for both feature sets ($\mathcal{F}$ and $\mathcal{F} \backslash \mathcal{F}_{pc}$) after applying feature selection. We observe that on the one hand SVM is able to leverage SFFS to extract meaningful subsets of features. On the other hand, SVM reveals dramatic performance drops if no feature selection is applied. At the same time, Perceptron does not outperform the distribution-based (DB) and the 0-R predictors, which are both just baseline predictors.

For the NST prediction task, we observer in Figure 1d, Figure 1e, and Figure 1f that all predictors achieve a high accuracy. It is not surprising since the class of "No Transitions" dominates with almost 95%. Only Perceptron fails for a number of users to achieve a high accuracy. In the case of the F1 score and MCC the predictors k-NN, CART, and SVM achieve similar performance after applying feature selection to both sets – with and without the phone context data. It is worth to note that for both metrics – F1 score and MCC – the predictor CART tends to be much more robust in the cases with no feature selection by achieving at least twice as high performance than the next best performing predictor.

Last but not least, Figure 1g, Figure 1h, and Figure 1i demonstrate results for the NP prediction task. For k-NN, CART, and SVM we make two observations. First, the feature selection leads to performance improvements in terms of all 3 metrics. Second, the consideration of the phone context data does not show any signification improvements in terms of the considered metrics.

4. CONCLUSIONS AND FUTURE WORK

We summarize our results and conclude that our use of the phone context features *does not* lead to signification performance improvements. However, phone context data is in general a rich information source. Statistical information from the data set highlights the presence of the potentially meaningful context data and the fact that the phones have been used intensively by their owners. Thus, we believe that instead of capturing temporal phone information, e.g., time since last action X, analysis on the correlation between the appearance of phone data events, e.g., received a phone call, and a corresponding mobility behavior may uncover additional potential for prediction improvements. In the context of this work, we leave the proof of this hypothesis as future work.

5. ACKNOWLEDGEMENTS

This work has been partially supported by the Collaborative Research Center 1053 funded by the German Research Foundation.

6. REFERENCES

[1] C. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[2] J. Krumm and D. Rouhana. Placer: Semantic Place Labels from Diary Data. In *UbiComp'13*, 2013.

[3] J. Laurila *et al.* The Mobile Data Challenge: Big Data for Mobile Computing Research. In *Pervasive'12*, 2012.

[4] P. Baumann *et al.* The Influence of Temporal and Spatial Features on the Performance of Next-place Prediction Algorithms. In *UbiComp'13*, 2013.

[5] V. Srinivasan *et al.* Mobileminer: Mining Your Frequent Patterns on Your Phone. In *UbiComp'14*, 2014.

[6] X. Wu *et al.* Top 10 Algorithms in Data Mining. *Knowledge and Information Systems*, 14(1):1–37, 2008.