Leveraging Periodicity in Human Mobility for Next Place Prediction

Bhaskar Prabhala, Jingjing Wang, Budhaditya Deb, Thomas La Porta, and Jiawei Han

Periodic transitions from place to place are inherent in human movements. Through visual examination we detect these periodic movements in traces of user tracking data. However such user tracking data sets tend to be sparse and incomplete. In addition, periodic movements are surrounded by noise: transitions to and from less frequently visited places and transitions to one of a kind visits. We present algorithms leveraging techniques and models to detect periodicity in individual user movements. Our algorithms predict a user's next place given only the current context of timestamp and location. We apply these algorithms to real user mobility data sets. Prediction accuracy depends on the ratio of periodic movements to noise in user traces. For a majority of users in a movement tracking data set collected over a year, our algorithms achieve next place prediction accuracies of 50% and above.

Algorithms

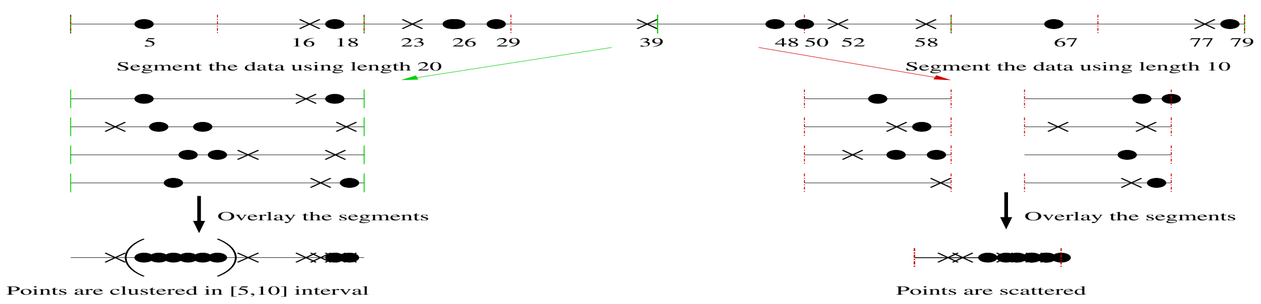

Fig. 3. Intuition of the Periodicity Model

- Observations form dense clusters along the timeline when we overlay with correct period
- PeriodicaS is a Multi-Class Classification Algorithm via SVM with RBF Kernel
- Created Training and Test data from the datasets

Algorithm 1: PeriodicaB - Periodicity Based Algorithm

Input: current place p_{test}, current time interval t_{test},threshold, profile of the user $\{distp_i \rightarrow p_j\}, p_i, p_j$ could be any places the user has visited

Output: next place $p_{predict}$

if *there exists a place* p_n, *s.t.* $confidence(p_{test} \rightarrow p_n) > threshold$ **then**
  $p_{predict} = p_n$;
else
  $p_{predict} = \arg\max_{p_j}(sim(dist(t_{test}), dist(p_{test} \rightarrow p_j)))$;
return $p_{predict}$;

Prediction Accuracy

Algorithm	*MDC*	*WTD*
PeriodicaB	46.78%	49.04%
PeriodicaS	48.93%	45.58%

TABLE I
AGGREGATED AVERAGE ACCURACY

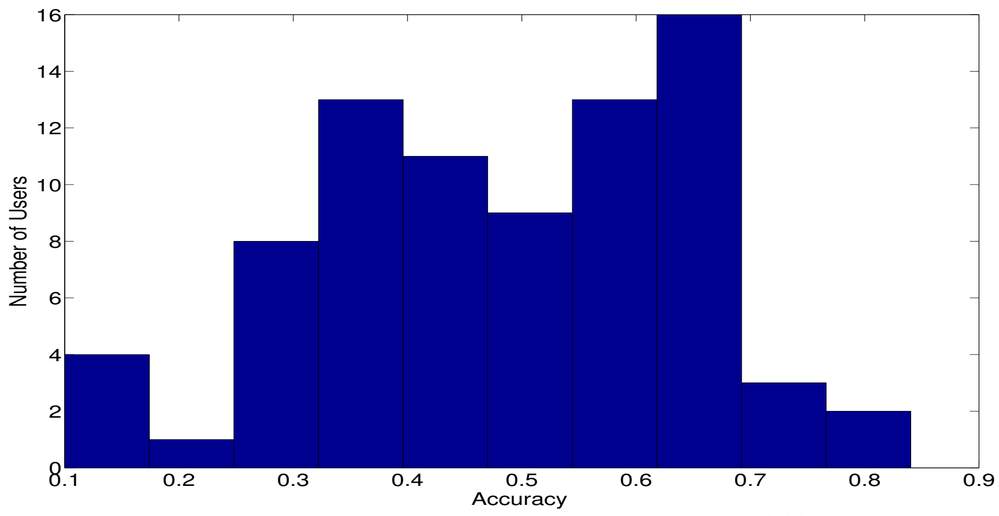

Fig. 4. Accuracy of predictions for users in MDC data set

Impact of Periodicity

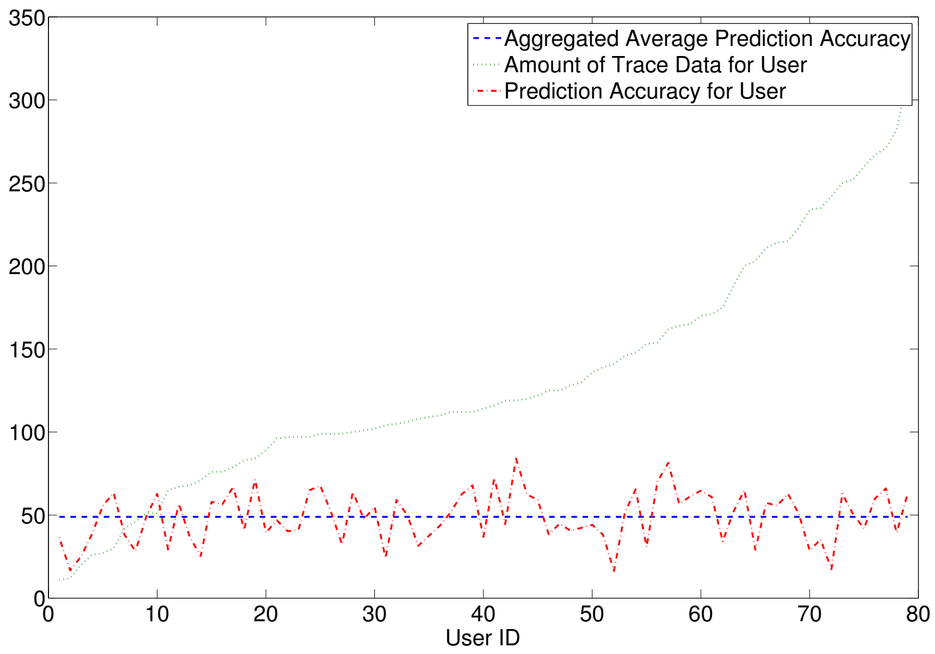

Fig. 6. Accuracy of predictions VS amount of trace data for each user in MDC data set

Summary

- We use MDC and WTD datasets
  - MDC data from ~200 Nokia smart phone users in Lake Geneva region (Switzerland)
  - WTD data from ~275 UCSD freshmen PDA users for an 11-week period during the Fall Semester 2002
- No place semantic information is available nor derived
- Periodicity in movements from office to lunch are lost if the lunch places are at different locations
- Periodicity is mined from the trace data with incomplete observations and noise
- Experiments were conducted in Matlab using the interfaces of LIBSVM
- Aggregated average accuracy is 50% whereas for majority of users in the MDC data set prediction accuracy is in the range of 40% to 70%

Related Publications

[1] J. K. Laurila, D. Gatica-Perez, I. Aad, J. Blom, O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, and M. Miettinen, "The mobile data challenge: Big data for mobile computing research," in *Proceedings of Mobile Data Challenge by Nokia Workshop, in conjunction with Int. Conf.. on Pervasive Computing,* Newcastle, June 2012.

[2] M. McNett and G. M. Voelker, "Access and mobility of wireless pda users," SIGMOBILE Mob. Comput. Commun. Rev., vol. 9, no. 2, pp. 40–55, Apr 2005. [Online]. Available: http://doi.acm.org/10.1145/1072989.1072995

[3] N. Eagle, A. S. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences,* vol. 106, no. 36, pp. 15274–15278, 2009. [Online]. Available: http://www.pnas.org/content/106/36/15274

[4] Z. Li, J. Wang, and J. Han, "Mining event periodicity from incomplete observations," in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, ser. KDD '12. New York, NY, USA: ACM, 2012, pp. 444–452. [Online]. Available: http://doi.acm.org/10.1145/2339530.2339604

Sponsored by U.S. Army Research Laboratory under the Network Science