Improving Grid-Based Location Prediction Algorithms

Itay Hazan, Dr. Asaf Shabtai

Department of Information Systems Engineering,
Ben-Gurion University of the Negev, Beer-Sheva, 84105, Israel
itayhaz@post.bgu.ac.il, shabtaia@bgu.ac.il

There is a vast research upon the problem of *location prediction* starting from the invention of the GPS and getting back attention due to the spread use of mobile devices containing GPS trackers. Existing methods for location prediction consists of two main parts: *Spatial division* that break the surface into smaller units and a *probability function* that maps a probability of moving from one unit to another according to both the previous movements of the user and contextual information gathered from the mobile device (e.g. day, hour or recent phone call).

There are two main types of spatial division approaches: *grid-based*, and point of interest *(POI)-based*. In grid-based approach the map splits to cells according to a predefined grid and the prediction is for the next cell on the grid in which the user is expected to visit. In the POI-based approach several point-of-interest locations are marked as interested or might be interested for the user through expert tagging or automatic extraction, while the function predicts the next POI of the user. In both methods, at the moment of the prediction, the probability function map probabilities to cells (also refers to POIs) which indicates the likelihood of the user to visit them in some time in the future. The vector of cells and their probabilities produced by the algorithm will be referred as the *probability vector.*

According to our knowledge location prediction algorithms consider previous routes, the current location of the user, day, hour and other contextual features that has to be learnt over time. While there is a constant tradeoff between the contextual features and the amount of data needed for the algorithm to produce confident predictions. If the algorithm considers more contextual features such as the day of the week, hour, and the current location it will take more time to gain good results but eventually it will gain better results from algorithm that does not consider the current location of the user or the day of the week.

We suggest that the direction of the user is not a contextual feature that has to be learnt from scratch, and can be used as a zero-day heuristic method to improve location prediction learning algorithm if it is correctly used and may even improve over time. Previous location methods tried to use solely the direction of the user in order to predict the location but they did not manage to achieve better results than learning algorithms. In this work we suggest a method that combines learning algorithms with the heuristic method of the direction along with a dynamic divide-and-conquer technique that splits the map into areas and transform the probabilities in the vector according to the areas and it's distribution. And we show how our method can improve the probability vector produced by the learning algorithm, thus, improving the prediction of the location and radius in up to 10% out of the total predictions.

We evaluated our proposed method using two datasets collected in the cities of Jerusalem and Beer Sheva and two well-known algorithms in the field of Grid-based location predictions. The first is Markov Model based and the second is Frequent Cells based as we will explain later in further. We used only Grid-based methods in the evaluations as inferring the type of location through POI is a challenging task and usually requires human assistance with continuous labeling of POIs. In both cases we received an accuracy improvement in which we affirmed through paired t-tests.

Another issue we will discuss is that multiple studies have attempted to predict the user’s next location, but to the best of our knowledge only a few of them dealt with the accuracy of predictions and tried to leverage the above factors in order to *dynamically* set a confidence area for the prediction.

We focus on the grid-based approach and propose several methods for providing a dynamic and effective confidence area for each predicted location. We define *confidence area* as the extended area around the predicted location that the user might be in, and *confidence* as the level of assurance that the user will be in the predicted cell/area. The proposed methods can be applied to the output of *any* grid-based location prediction algorithm and do not depend on the configuration of the grid or the maximization functions that determine which cell to predict.

We evaluated the proposed methods using two well-known datasets in the field of location prediction and also used a third dataset we collected ourselves in the city of Beer Sheva. Our results demonstrate the effectiveness of our proposed methods. Using our methods we were able to increase the accuracy of two well-known location prediction methods by 5 to 15 percent while using, on average, smaller confidence areas when compared to the fixed radius approach used in previous studies. The proposed methods also reduce the incidence of mistakes resulting from the natural error of mobile device location sensors and different size of locations by inspecting the surrounding cells and taking previous errors into account.

References

1. DO, Trinh Minh Tri; GATICA-PEREZ, Daniel. Where and what: Using smartphones to predict next locations and applications in daily life. *Pervasive and Mobile Computing*, 2014, 12: 79-91.
2. De Domenico, Manlio, Antonio Lima, and Mirco Musolesi. "Interdependence and predictability of human mobility and social interactions." *Pervasive and Mobile Computing* 9.6 (2013): 798-807.
3. DO, Trinh Minh Tri; GATICA-PEREZ, Daniel. Contextual conditional models for smartphone-based human mobility prediction. In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, 2012. p. 163-172.
4. Krumm, John, and Eric Horvitz. "Predestination: Inferring destinations from partial trajectories." *UbiComp 2006: Ubiquitous Computing*. Springer Berlin Heidelberg, 2006. 243-260.
5. Song, Chaoming, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. "Limits of predictability in human mobility." *Science* 327, no. 5968 (2010): 1018-1021.
6. Chon, Yohan, Hyojeong Shin, Elmurod Talipov, and Hojung Cha. "Evaluating mobility models for temporal prediction with high-granularity mobility data." In*Pervasive computing and communications (percom), 2012 ieee international conference on*, pp. 206-212. IEEE, 2012.
7. Lu, Xin, Erik Wetter, Nita Bharti, Andrew J. Tatem, and Linus Bengtsson. "Approaching the limit of predictability in human mobility." *Scientific reports* 3 (2013).