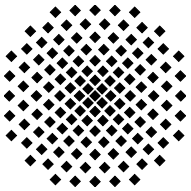

Universität Stuttgart

Institute of Parallele and
Distributed Systems (IPVS)

Universitätsstraße 38
70569 Stuttgart

M. Sc. Zohaib Riaz
+49 711 685 88296
zohaib.riaz@ipvs.uni-stuttgart.de

Masters Thesis

Design and Development of Software Agents for Location Privacy-risk estimation

Student:
Tutor: M. Sc. Zohaib Riaz
Examiner: Prof. Dr. Kurt Rothermel
Timeframe: 6 months starting anytime

1 Background

The use of Location Based Applications (LBAs) has become prevalent with the fast adoption of location-enabled devices such as smart phones. A typical example of these applications are geo-social networks (e.g., Foursquare, Google+, Facebook, etc.) where users share their location data, along with other content, with their family and friends. However, sharing personal location data with *non-trusted* service providers as well as malicious acquaintances can be detrimental to a user's privacy. For example, location information can reveal a person's habits, their inclinations (political, religious etc.), as well as other sensitive information such as the state of their health. Among other reasons, such privacy leaks may also be attributed to the difficulty faced by LBA users in judging the *privacy-risk* associated with location information sharing.

To help LBA users in anticipating this privacy-risk, intelligent software is needed that can estimate the loss of user privacy as a result of location sharing. As an example, whenever a user wishes to share his location with a social connection, a software agent running on his mobile device estimates the privacy-risk based on the current user location as well as his past location sharing activity and pushes this information to the user as a notification. In view of the information in the notification, the user can then decide for or against sharing his current location. Obviously, the actual amount of privacy-risk may depend on various factors, such as the background knowledge of the target social connection about user location, making its accurate estimation non-trivial. However, we can reasonably limit the scope of this estimation by defining privacy-risk as the increase in predictability of user's next location as a result of sharing their current location. As a simple example, before sharing his current location, say *office*, at 9 am, Bob receives a privacy-risk notification (from the software agent) that sharing the current location makes his future location (also *office*) highly predictable for the next 8 hours (his usual working time). Hence Bob can use this auxiliary information to decide whether he really wishes to inform others about his current location or not.

2 Goals and Tasks

To achieve the above functionality, the goal of this thesis is to design and evaluate software agents for privacy-risk estimation. The basic functionality of an agent is as follows. The agent regularly collects the location information of the user and also keeps a record of when and how often does he share his location with various third parties, i.e., LBAs and their social connections. Based on this information, the software agent builds and maintains *next-location prediction models* for each third party. When the user wishes to share location with party x, the software agent can use the *next-location prediction model* for party x to predict the impact of sharing the current location, as discussed in the previous example, and displays this information to the user.

The detailed tasks in the thesis are the following:

- Getting familiar with research work on next-location prediction (e.g.[4, 2, 1, 3]).
- Implementation of various next-location prediction algorithms.
- Evaluation of the above implemented algorithms with regard to their accuracy of privacy-risk estimation on real-world location datasets.
- Development of an Android application that integrates high accuracy next-place prediction algorithms (as determined in the above evaluation step) to represent the software agent.
- *Documentation* of the results in a written report.
- *Presentation* of the results in a final talk in the colloquium of the Distributed Systems Department.

3 Requirements

The thesis demands good programming skills (particularly in Java for implementing the Google Android application) as well as good understanding of the basic probability theory (for developing next-location prediction models).

4 Grading

Grading will be done based on the quality of the conceptual and practical results, the quality of the documentation in the written report, the conceptual and practical knowledge of the student (also gained during the work), and the general procedure according to best practices.

References

[1] P. Baumann, W. Kleiminger, and S. Santini. The influence of temporal and spatial features on the performance of next-place prediction algorithms. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '13, pages 449–458, New York, NY, USA, 2013. ACM.

[2] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez. Next place prediction using mobility markov chains. In *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*, page 3. ACM, 2012.

[3] J. B. Gomes, C. Phua, and S. Krishnaswamy. Where will you go? mobile data mining for next place prediction. In *International Conference on Data Warehousing and Knowledge Discovery*, pages 146–158. Springer, 2013.

[4] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo. Mining user mobility features for next place prediction in location-based services. In *Data mining (ICDM), 2012 IEEE 12th international conference on*, pages 1038–1043. IEEE, 2012.