論 文 の 内 容 の 要 旨

論文題目　Study on Crowd Sensing for Analyzing Human Activities and Urban Environment
（人間活動と都市環境分析のためのクラウドセンシングの研究）

氏　　名　　劉　広文

Pervasive smartphones that embed a variety of sensors enable us to sense and learn about not only the physical environment around us but also the society we live in. On the other hand, crowd sensing is a new sensing approach that individuals with sensing and computing devices collectively share data sensed or generated from their mobile devices, and aggregates the data in the cloud for discovering knowledge and solving problems. We designed a sensing platform based on crowd sensing paradigm to inexpensively acquire sensor data from smartphones, such as ambient noise, light, location, acceleration, temperature, etc., at a large scale. Based on the sensor data collected in our sensing experiments, we mainly studied trajectory computation, sensor data management, sensor application, and data analysis for specific purpose.

Trajectory obtained from GPS is a key factor when analyzing human mobility and transport issues. Trajectory simplification can greatly improve the efficiency of data analysis (e.g. trajectory querying, trajectory clustering). Based on the observation of information content contained by sampling data, we assume that (1) the sampling points on the boundary of MBR (Minimum Bounding Rectangle) contain more information content, (2) the bigger the area of MBR is, the more the points should be stored. We applied these two assumptions in our method to simplify trajectory online. Two main components of this method (i.e. divide/merge principle and selection strategy), are elaborated in the thesis. Moreover, we define a new error metric - enclosed area metric - to evaluate the accuracy of simplified trajectories, which is proven more robust against the uncertainty of GPS. Through comparing with other methods in a series of experiments over huge dataset, our method is proven effective and efficient. Furthermore, our method is a pure online procedure which can be readily installed at the mobile terminal to preprocess trajectory before sending it to back-end server.

We also solved another major issue: the sheer volume of data collected through crowd sensing can deeply hamper the performance of various applications, for example, storage, transmission, power consumption and data processing cost. We proposed a method to reduce the volume of sensor data while preserving the information content of the original data. We rethink data reduction problem by analogy to electoral system from the viewpoint of data diversity. Hence, after data reduction, output data (target) can represent the original data (source) as parliament members are elected to delegate their constituencies. This method can compress multi-dimensional data without any assumption on distribution. Although we didn't theoretically prove our model is effective, we draw our conclusion based on a large amount of experiments with real data set. Through comparing with conventional methods and state-of-the-art methods, our method performs well in terms of data divergence especially Itakura-Saito distance. Our method also outperforms other methods in terms of clustering performance. It is foreseeable that data reduction will linearly mitigate storing and transmitting problems, though we did not evaluate it by experiments.

Furthermore, we developed a practical system (called iBaro-altimeter) to measure elevation by using smartphone's barometer. Accurate estimation of elevation is important for many location based services, such as 3D navigation, floor localization, emergency rescue, etc. Although, it is possible to obtain altitude from GPS, its accuracy is unreliable and applicable in outdoors only. It is possible to use barometers on smartphones to estimate elevation in both indoor and outdoor scenarios. However, to reach an acceptable level of accuracy, a “reference point” which periodically broadcasts its air pressure and temperature is required. We proposed a method to increase the spatio-temporal density of reference points by exploiting neighboring smartphones as ad-hoc reference points. In addition, we also employed Kalman filter to stabilize the elevation profile due to the instability from smartphone sensors. To the best of our knowledge, this is the first time to measure elevation (absolute height) with high accuracy using smart-phone. A series of experiments conducted in both indoor and outdoor environments with varying geographic characteristics using different models of smartphones, reveal that the errors are less than 3 meters in outdoor walking, 6 meters in mountain climbing, 0.9 meter in indoor floor localization. These errors are acceptable for most practical applications. For instance, in case of indoor floor localization, even with our recorded error of 0.9 m it is still possible to reliably detect floors.

To explore collective intelligence is the main goal in crowd sensing. By aggregating sensor data from a group of participants, we can learn about human activities and urban environment. Last but not least, we analyzed sensor data to discover knowledge from crowd intelligence for specific purposes: (1) estimate how busy an area is by using ambient light and noise from smartphone sensors; (2) build noise map for a residential area to find noisy area; (3) sense and estimate micro-scale weather by integrating smarphone's weather sensors with meteorological stations.

There is a lot of work left for us, since crowd sensing is still at its infancy, for instance, robust frame work, privacy preservation, incentive mechanism, and large scale useful applications. In future, we believe that new smart wearable devices, such as, smart glasses, smart watches and smart clothes, will make people more easily involve into sensing, which will boost the development of crowd sensing.