Expose Urban Activities from Human Flow

Teerayut Horanont[1], Apichon Witayangkurn[1], Yoshihide Sekimoto[2] and Ryosuke Shibasaki[2]

[1]Institute of Industrial Science, the University of Tokyo, Japan
[2]Center for Spatial Information Science, the University of Tokyo,
{teerayut, apichon}@iis.u-tokyo.ac.jp, {sekimoto, shiba}@csis.u-tokyo.ac.jp

Abstract

Today, the urban computing scenario is emerging as a concept where humans can be used as a component to probe city dynamics. The urban activities can be described by the close integration of ICT devices and humans. In the quest for creating sustainable livable cities, the deep understanding of urban mobility is of crucial importance. This research aims to explore and demonstrate the vast potential of using large mobile GPS datasets for the analysis of human activity and urban connectivity. The new type of mobile sensing data called "Auto-GPS" has been anonymously collected from 1.5 million people for a period of one year in Japan. The analysis delivers some insights on interim evolution of population density, urban connectivity, events and diversity of urban functions. The results enable planners to better understand the urban organism with more complete inclusion of people's movement and their evolution through space and time.

1. Introduction

New technology can help cities manage guarantee and deliver a sustainable future. In the past few years, it has become possible to explicitly represent and account for time-space evolution of the entire city organism. Information and communication technology (ICT) has the unique capability of being able to capture the ever-increasing amounts of information generated in the world around us, especially the longitudinal information that enables us to investigate patterns of human mobility over time. Thus, the use of real-time information to manage and operate the city is no longer just an interesting experience but a viable alternative for future urban development.

In this research, the analysis of mobile phone location, namely "Auto-GPS", has been used to serve as frameworks for the variety of measures of effective city planning. More specifically, we explore the use of location information from Auto-GPS to characterize human mobility in two major aspects. First is the commuting statistics and second is the city activity, how the change of activities in part of urban space can be detected over times.

In general, a classic travel survey is frequently used to acquire urban connectivity and trip statistics. However, they truly lack of long-term observation and sample size is always the main limitation due to the highly cost and extra processing time. In this paper, we propose a novel approach that takes advantage of anonymous long-term and preciously collected spatial-temporal location generated by Auto-GPS function from ordinary mobile phone users. As of the best of our knowledge, this is the first time that large-scale GPS traces from more than one million users have been observed and analyzed countrywide for travel behavior research.

To have evidence showing clearly how this would help planning and decision making, we selected one of the major active area in central Tokyo called "Odaiba" as our study area. Odaiba is a large artificial island in Tokyo Bay, Japan. It was initially built for defensive purposes in the 1850s, dramatically expanded during the late 20th century as a seaport district, and has developed since the 1990s as a major commercial, residential and leisure area. Odaiba is suitable for this analysis since it is isolated from other parts of Tokyo. It provides all urban amenities like a small city including hotels, department stores, parks, museums, office buildings and residential areas.

The rest of the paper is organized as follows: section 2 outlines related work; Section 3 describes the dataset and the basics of Auto-GPS; Section

4 methodology; Section 5 explains the results from our analysis; and Section 6 provides conclusions and recommendations.

2. Related work

Location traces from mobile devices have been increasingly used to study human mobility, which is important for urban planning and traffic engineering. Several aspects of human mobility have been exploited. Human trajectories show a high degree of temporal and spatial regularity with a significant likelihood of returning to a few highly visited locations [1]. Despite the differences in travel patterns, there is a strong regularity in our mobility on a regular basis, which makes 93% of our whereabouts predictable [2]. Understanding mobility patterns would yield insights into a variety of important social issues, such as the environmental impact of daily commutes [3]. It has been reported also the potential use of mobile positioning for the development of real-time monitoring tools for tourism planning and management [4].

These recent studies have emphasized on modeling, prediction, and interurban analysis of human mobility, but not on the richer context of it such as the engaged activity in the location visited. There are many studies that use GPS records to identify trip trajectories. Most of these works begin with the segmentation of GPS logs into individual trips, usually when there is a significant drop in speed [5][8], or when GPS logs remain in one area for a certain amount of time [6][7].

With the advance of today's ICT technologies, it is possible to realize a sort of socio-technical superorganism to support high levels of collective "urban" intelligence and various forms of collective actions [8]. It therefore becomes our interest in this work to investigate on how to use large-scale, long-term GPS data from mobile phones to extract valuable urban statistics and to project the real world information.

3. Dataset

There were two datasets used in this study. The main dataset was collected from approximately 1.5 million mobile Auto-GPS users of a certain mobile phone service provided by a leading mobile phone operator in Japan. Under this service, handsets provide a regular stream of highly accurate location data, and thereby enable support services that are closely linked with the user's behavior. Technically, an Auto-GPS-enabled handset position is measured within 5 min and sent through a network of registered services. (Fig.1) The data was recorded from August 2010 to October 2011.

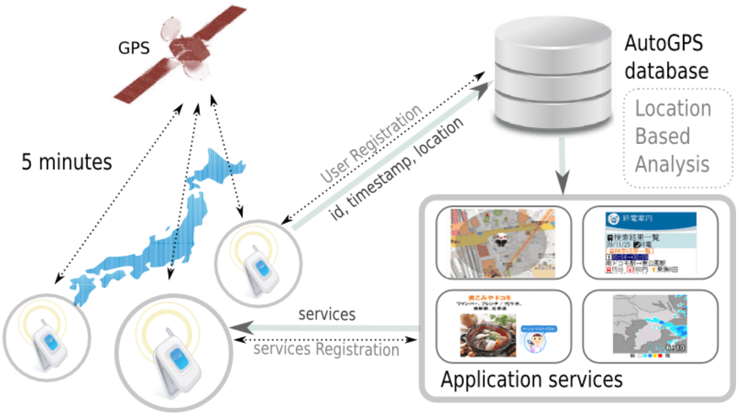

Fig. 1. Flow diagram illustrating the Auto-GPS services supported by Japanese handsets since 2009

In order to preserve user privacy, Auto-GPS data is provided in a completely anonymous form to ensure privacy of personal information. It is important to acknowledge that there is some selection bias in this dataset, as participants are limited to users of a specific mobile phone service. The distribution of user type was estimated from 50,000 online surveys and about 2.6 percent (1,356 respondents) replied to use this service. In addition, this service uses adaptive techniques to minimize battery drain by only activating GPS when the accelerometer moves in a recognized way. This results to an average of 37 GPS points per user per day. Fig. 2 shows a graph of the average number of GPS points per day in this dataset. A small sample of the raw data is shown in Fig 3.

The second dataset is the census data used for validation purpose. The census data was released in 2008, which represents census data for a grid size of 1 by 1 km. This data was provided by the National-Land Information Office. [10]

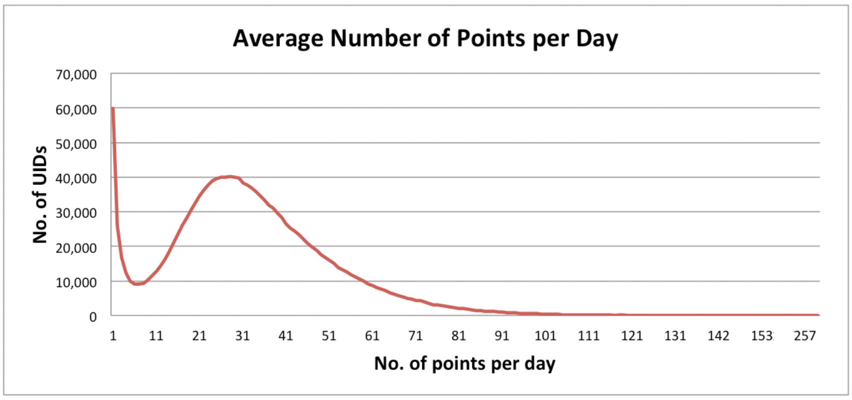

Fig. 2. The average number of GPS points per day is 37, indicating that Japanese users spent approximately 3 hours traveling with their handset each day.

Dummy-ID	Time	Latitude	Longitude	Error	Altitude
00862690	2010-08-01 12:01:09	34.69888	135.534146	1	64.00
00862754	2010-08-01 21:10:13	39.703028	141.146445	2	176.94
00886354	2010-08-01 12:48:23	34.33872	135.600167	3	165.73
00862690	2010-08-01 14:46:09	34.709877	135.591781	1	64.00
00169966	2010-08-01 18:19:52	35.534478	140.304336	3	39.64
00169966	2010-08-01 18:24:52	35.527892	140.312319	3	17.83

Fig. 3. A sample of Auto-GPS data that include an anonymous dummy-id, timestamp, geo-location, error level and altitude. The error level indicates the strength of the GPS signal available to the handset.

4. Methodology

Before we can do the analysis part, it is necessary to work on the pre-processing and data mining steps. This data is considered as big data, having a total of 9.2 billions GPS records. We processed 1 year of Auto-GPS data to calculate the number of people who live within the specific area, more specifically, finding their home location. First, we extracted the locations of daily rest, or “stay points,” for each individual trajectory. Let P represent a set of sequential traces of the user such that $P = \{p(1), p(2), p(3),\dots, p(i),\dots\}$ where $p(i)$ is the i^{th} location of the user and $p(i) = \{\text{id}, \text{time}, \text{lat}, \text{lon}\}$. A stay point can be identified as a series of locations in which the user remains in a certain area for a sufficiently period of time, where distance in space and difference in time between observed points are applied as constrained multicriteria in the detecting method:

$$\text{Distance}(p_{start}, p_{end}) < D_{threh} \text{ and TimeDiff } (p_{start}, p_{end}) > T_{threh}$$

where D_{threh} and T_{threh} are adjustable parameters. D_{threh} is the maximum coverage of movement for which an area can be considered a stay point, and T_{threh} is the minimum time the user must spend in the stay point. For our purposes, a stay point is detected if $D_{threh} <= 150$ meters and $T_{threh} > 20$ minutes. The most typical stay points from midnight to 6 a.m. are assumed to be home locations. This yielded a fairly accurate estimation of home locations, comparable ($R^2 = 0.79$) to the population density recorded in 2008 census data provided by the Statistics Bureau (Fig. 4). The horizontal axis is the number of population count from the census data in 1-kilometer grid and the vertical axis is the estimated population at the same grid using our home location extraction algorithm.

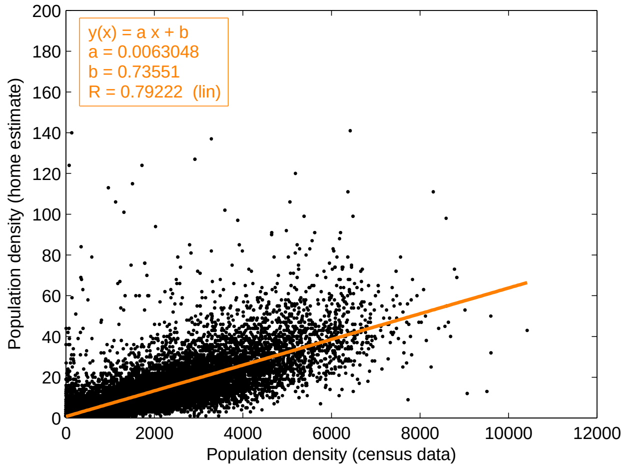

Fig. 4. Comparison of estimated home locations from Auto-GPS and to census data per 1-kilometer grid section.

From the above result, we considered our “stay points” are reliable to do further analysis. The stay points and home locations were used as inputs for calculation of various urban indicators and statistics across different spatial and temporal levels in the next section.

5. Results and Discussions

Finding the urban descriptive knowledge such as where/how/when/why of the people who use the area is one of the most important information for urban planners. Our first result attempted to explain where the people come from. We constructed multiple criteria to define visitors in the area. We used the minimum stay of 30 minutes and excluded people who have home and work location in the area. The maximum annual visit is set to 8 times as it is the third quartile of the entire dataset. (Fig. 6) The annual total of visitors to Odaiba area was estimated at 80,463 people from 1.5 million total samples or 5.36% of the population. Fig. 5 shows the choropleth map of estimate annual visitors. As expected, the nearer the prefecture is to the Odaiba area, the more visitors are coming from there. There is an exception for the big city such as Nagoya, Osaka, Fukuoka and Hokkaido where air transport services are operated more often.

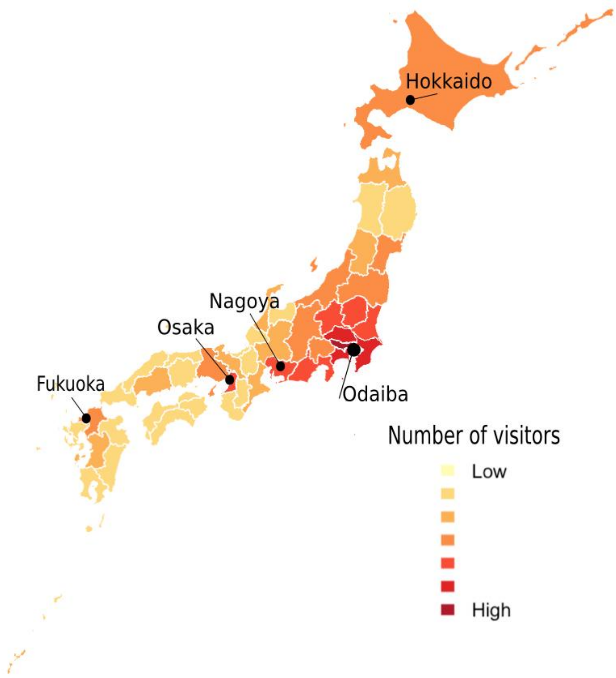

Fig. 5. Estimated annual visitors to the Odaiba area by prefecture.

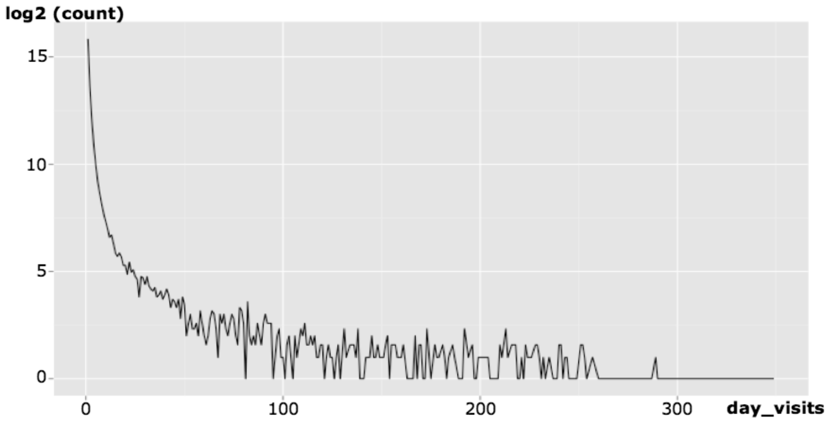

Fig. 6. Count of the number of day visits to the Odaiba area.

When observing the number of visitors in Odaiba per day (see Fig. 7), it appears that the area is more popular during summer. This is because of the big event arranged by Fuji TV. The highest number of visits to Odaiba was on August 14 when the Tokyo Bay Grand Fireworks Festival was held. The second most visits were on the Christmas Eve, as the area is known as a couples place. We notice a significant distinct drop in the number of visitors suddenly in March 11. It was the day when an earthquake of magnitude 9.0 hit Japan in 2011, followed with the "radiation leakage" of 2 nuclear power plants in the Fukushima prefecture. We could observe 3 weeks of an abnormal reduction in visiting the area before it returned to the normal situation.

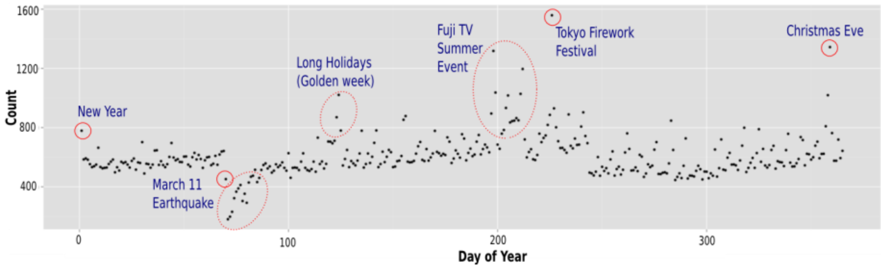

Fig. 7. Estimated daily visits to Odaiba area. The magnitude of anomalies can vary greatly between events, and this could lead to composite dominated by a few major events.

Next, we visualized how different the activities between weekdays and weekend are by overlaying weekdays stay points over the weekend. (Fig. 8) Surprisingly, there are several clusters that highly dominate over others at particular locations. By incorporating prior knowledge of the area and collection of news, it revealed clear evidence how the patterns are created. The locations marked with “a” are complex buildings where shopping malls, hotels and restaurants are located in. This yields the similar distribution of both weekdays and weekend. The “b” mark is an open space where we can see half of the areas are more active during the weekend. This is because of the special events are usually held only on the weekend. The area in the upper part is served as outdoor parking space and is the main area of Fuji TV summer event. This event is usually held for 3 months in summer regardless weekend or weekday. The “c” areas are event spaces that are mainly occupied during weekend. The “d” areas are office building that is more dominant at weekdays. Please note that the d1 area is the construction area during our data collection period.

By accumulating annual statistics of visitors, Fig. 9 provides valuable population count information in each building for the entire year. The height of the buildings corresponds to the number of visitors. It is clear that Shopping malls, restaurant and complex type buildings are the most destination place in Odaiba area. All of them have 10 times more visitors than the office area.

6. Conclusions

This study explores the potential of using mobile Auto-GPS enables in the new context and broader advances towards the understanding of today’s excessive mobility. The finding of this remarkable dataset is to capture the urban evolution from the real movement of people. The results display summarizes the findings of the comprehensive and creative process using Auto-GPS data. Compared to existing methods, Auto-GPS contains added values that providing accurate and valid information of quantitative spatio-temporal data. It introduces a new framework for urban analysis driven by the entire population of the city. Finally, the importance of this research ultimately lies on how it can be practically applied and utilized massive amount of high accurate GPS data for future sustainable urban development.

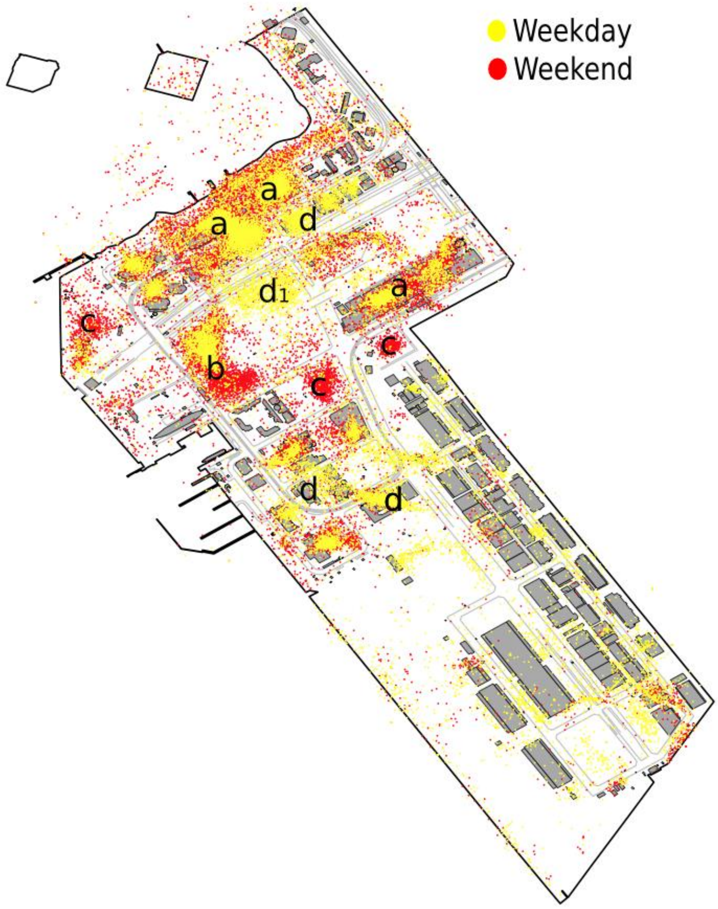

Fig. 8. Comparison of estimated people activities between weekdays (yellow) and weekend (red). The functions of built environment and urban patterns can be clearly depicted on map by considering the temporal aspect.

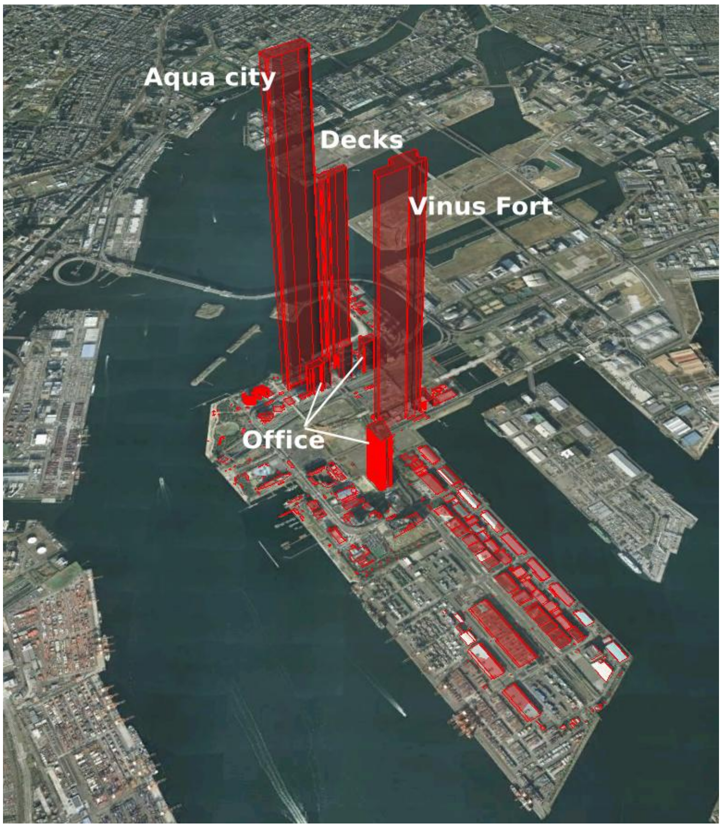

Fig. 9. The annual visits counted in every buildings of the Odaiba area. The heights of buildings represent the number of visitors who stop by.

Acknowledgements

This research carried out under the support of DIAS (Data Integration and Analysis System) by MEXT (Ministry of Education, Culture, Sports, Science and Technology). We would like to thanks ZENRIN Data Com (ZDC) CO., LTD. for archiving the dataset used in this research.

References

[1] M. C. Gonza´lez, C. A. Hidalgo, and A.-L. Baraba´si, “Un- derstanding individual human mobility patterns,” Nature, vol. 458, pp. 238–238, 2009.

[2] C. Song, Z. Qu, N. Blumm, and A.-L. Baraba´si, “Limits of predictability in human mobility,” Science, vol. 327, no. 5968, pp. 1018–1021, 2010.

[3] Isaacman, S., Becker, R., Cáceres, R., Kobourov, S., Rowland, J., & Varshavsky, A. (2010). A tale of two cities. Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications - HotMobile ’10, 19.

[4] R. Ahas, A. Aasa, S. Silm, M. Tiru, “Mobile Positioning Data in Tourism Studies and Monitoring: Case Study in Tartu, Estonia”, Information and Communication Technologies in Tourism 2007, pp. 119-128, 2007

[5] Robert de Jong and Wytse Mensonides. (2003). Wearable GPS device as a .data collection method for travel research. Working Paper, ITS-WP-03-02, Institute of Transport Studies, University of Sydney

[6] Hongmian Gong, Cynthia Chen, Evan Bialostozky, and Catherine Lawson. (2011). A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems*

[7] Ashbrook D, Starner T (2003) Using GPS to learn significant locations and predict movement across multiple users. Personal and Ubiquitous Computing 7:275–286.

[8] Li, Z., Wang, J., & Han, J. (2012). Mining event periodicity from incomplete observations. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’12, 444.

[9] Fontana, D., & Zambonelli, F. (2012). Towards an Infrastructure for Urban Superorganisms: Challenges and Architecture. The IEEE International Conference on Cyber, Physical and Social Computing.

[10] National-Land Information Office. Retrieved January 27, 2013, from http://www.mlit.go.jp/kokudoseisaku/gis/index.html