

Research on the Behavior Data Collection based on Mobile Location

He Zhang^{1, a}, Lanxin Hu¹, Zhaoming Zheng¹, Yi Tang¹, Zun Wang¹

¹Telecommunications engineering, SHANGHAI DIANJI UNIVERSITY, Shanghai, 201306, China

^aemail: 402226785@qq.com

Keywords: Location based services, Mobile data, Data collection

Abstract. With the popularity of mobile terminals and the mobile phone, Collecting individual, the collective spatio-temporal data have a lot of new technology. Corresponding also produced many terminal application of location-based services, they are constantly changing our lives, provide convenience for our life. Location-based services compared with questionnaire survey, the student survey and traditional mobile data traffic census method the features include contains the following features: (1) true information;(2) High precision time;(3) Save workload and staff;(4) suitable for large-scale data collection and analysis, reducing the loss of individual information; We select a campus, with student's real-time location data as samples, to introduce some data analysis methods.

Introduction

Research on the behavior data collection based on mobile location is an important branch of the research field in data mining, its significance lies in the implicit information through independent data mining and provides useful guidance to Community in need, such as Location Based services (Location -based service hereinafter referred to as the LBS) which is based on the user's Location information in mobile positioning system. According to the user's location information, it will determine the statistical probability of a certain behavior and provide the corresponding data analysis, finally giving service to meet the needs of users. In recent years, this kind of research has got more and more attention from all walks of life, also promoting its research and development of data mining technology.

LBS develops a number of commercial services, including information navigation system which allows the user to find the recent hotel and space distributed application of chat. According to different service object, the LBS services can be divided into individual users service, community service and third party service. (1) the individual customer service is the most simple form, LBS can provide both real and virtual space navigation services, such as driving navigation, tourist information accessibility, geographic yellow pages service, etc; (2) the group customer service will produce more applications, such as distributed chat, friends track, the game based on the location, Real-time traffic information, etc; (3) the third party service is the origin of the LBS services, it also the most potential applications. Third-party services can be collected by means of an anonymous summary the location of the mobile phone users and mobile data, spatial behavior model for urban studies; It also can be mobile information that is associated with users' personal information, which is a powerful tool for understanding and control the city phenomenon. It brings the opportunity and it is obvious to help life. For example, mobile banking, e-commerce, first aid, tracking service, and so on. Many softwares like, tencent, taobao, baidu map are obtained tremendous benefits from the LBS.

At the core of the LBS problem is to get the data and analysis data, the application is different, different applications need different data support direction. For example, taobao use LBS purpose is to help people shopping, food, travel, recommend different shopping menu, advertising and so on. Baidu map in order to facilitate people to travel. recommend different shopping menu, advertising and so on. In this paper, we study is aimed at students' this special group of LBS.

Data collection and application

Our test by getting 50 is willing to cooperate with the investigation of college students for a week in Shanghai motor real-time mobile phone signal data, and dynamic map carving. In the test of time, every time they occur a left area of the displacement, then send a location data collection personnel more variation signal. Instance (data collection as shown in figure 1, the designated areas as shown in figure 2, the sending location information as shown in figure 3) through our different location of mobile phone signal strength of each time period, the mobility characteristics of this part study groups.

Electronic Information Communication Engineering
Liu students

Monday 7:00 fifth dormitory 9:30 Canteen 9:55 Teaching Building 1 11:38 Canteen 12:10 Bathroom 12:25 fifth dormitory
14:00 Teaching Building 16:00 fifth dormitory 17:30 Canteen 17:57 fifth dormitory 23:00 Location termination

Tuesday 8:28 fifth dormitory 12:03 Canteen 12:22 Technology Center 14:13 fifth dormitory 17:36 Canteen 17:59 fifth dormitory
19:43 Bathhouse 20:25 fifth dormitory 23:00 Location termination

Figure 1 data collected examples

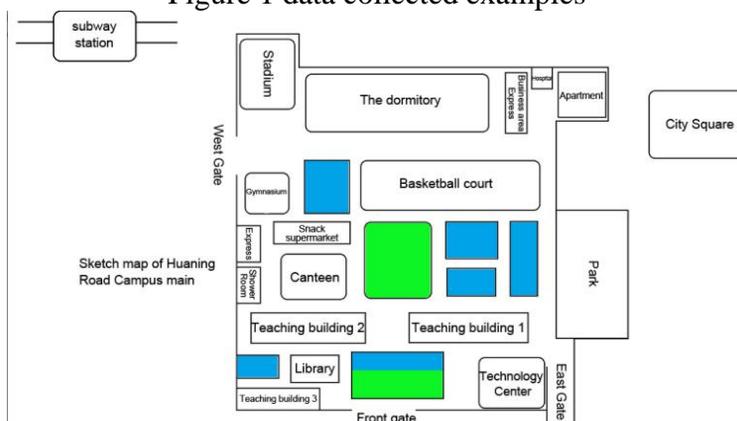


Figure 2 observation area map



Figure 3 specific location

By about 50 students each 7 days total 350 days of the data analysis, we have the following four periods to draw out of the cell phone use intensity distribution (with deeper color signal is larger number). (as shown in the figure 4, figure 5 and figure 6 and figure 7)

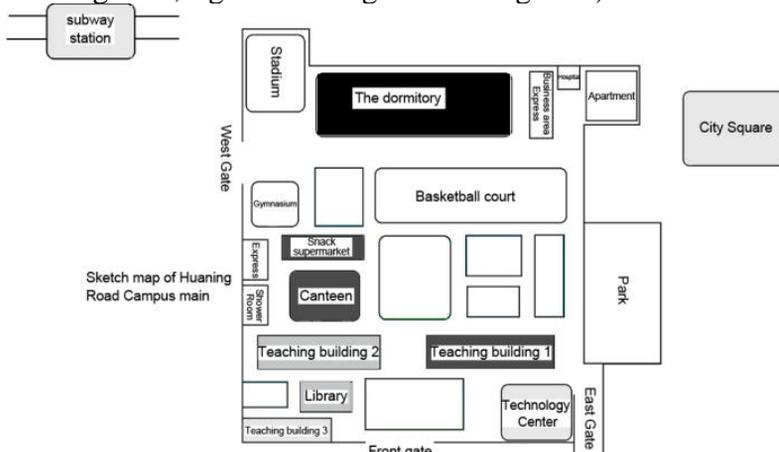


Figure 4 9:00-13:00 diagrammatic sketch of signal strength

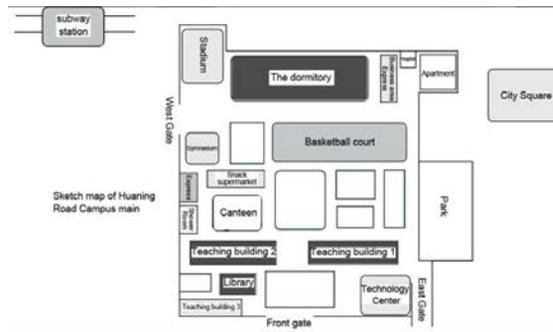


Figure 5 13:00-17:00 diagrammatic sketch of signal strength

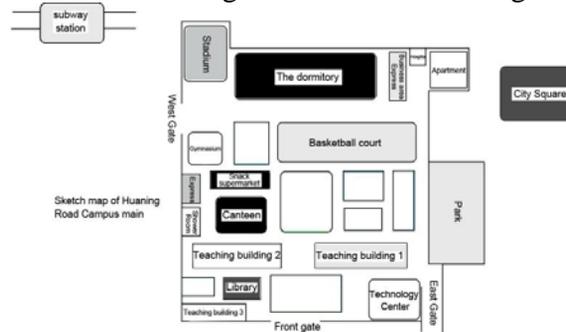


Figure 6 17:00-21:30 diagrammatic sketch of signal strength

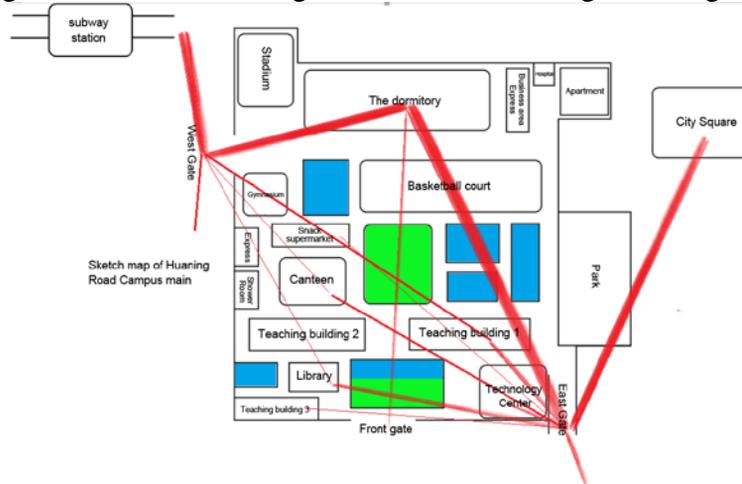


Figure 7 each school traffic statistical figure (Size of line thicknesses said traffic)

We can conclude from these illustrations, some possible problems in reality information to the group: 1, the dorm life occupied most of their time. 2, only less number of people has the habit of physical exercise.3, Courier transceiver, on the edge of a dormitory building in the utilization of relative bathroom side is low, but is closer. Online shopping is sending and receiving express, the number of the group of its shopping rate larger 4, covers an area of big park is almost no one concern.5, the traffic was heavy usage of the dormitory to the east gate line, the distance is relatively far, etc (as shown in figure 7).Due to the construction of campus has not yet been city square, station, and now with these sites, we think that this is the main cause of the school gate set unreasonable.

The semantic unknown area

The activity record of the people can be expressed by the track, if only research the path, We can not get completely and accurately the travel law of the people. We can related with environmental semantics, Digging deeper users travel patterns. Therefore, the following will study how to carry out the semantic expression of the trajectory.

Reside points, is a series of discrete anchor point, those discrete anchor point into the line, can be used to indicate the moving track of people's travel. Mobile trajectory data is different from other's data, it is a kind of spatiotemporal data, has its own features, these features recorded to

provide a path of travel. By extracting the trajectory residues, we can provide a trajectory is outstanding travel starting point and end point, and we can get the people's position that we are interested.

According to the frequency of students and the probability as well as the geographical position is different, we can put the campus-and-campus around different places is divided into five categories, shown in table 3-1.

Table 3-1 the semantic geographical position

number	Semantic category	keywords
1	live	The dormitory
2	diet	The canteen, snack street
3	learning,	Classroom、 library、 laboratory
4	entertainment	poly cinema; KTV; gym
5	shopping	Ouchan; BiJiang plaza; taobao, etc

In order to dig out the useful travel patterns, we choose one day as a period of time, the data is divided into units on track, and focused on observing the user travel more time, so you can observe the travel trajectory continuously. It is far from enough to dig people's characteristics from the trail itself, so there is a kind of feasible solution is to give the positioning data a certain semantic information.

Generally speaking, the rules of the travel not only law in space and also related with time, and is also the result of any interest. Space time rule describes the travel behavior, interested contact describes the correlation.

Describe the behavior of similarity has the following three types:

1. Completely similar: travel direction, distance, duration, track geometry etc.
2. The space synchronization: different trajectories have the same position. For example, sequence synchronization is people arrived in the same position in the same order; The sequence synchronization is likely to come to the same position in a different order.
3. Time synchronization: complete synchronization, for example, at the same time travel characteristics similar changes; Lag synchronization, after the lag time Δt travel have similar characteristics.

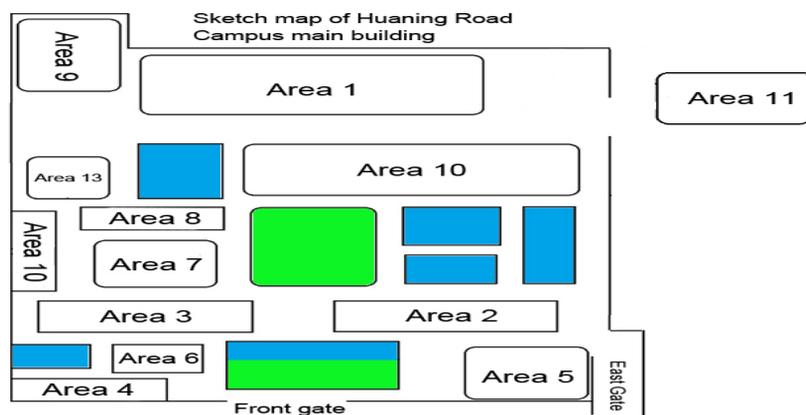


Figure 8 Campus (Jiangchuan Road NO.690)

We will through the analysis of the collected data, table 3-1 to figure 8. Data can be divided into working days and rest days, working days for the Monday to Friday, a day off for Saturday and Sunday. Travel through the analysis of the data of working days we found from the starting point to the finish line in the trajectory, (1) the 6:30 a.m to 7:30 a.m and after 9:30 p.m in the two periods, almost more than ninety percent of people in one area, thus we can infer that one area of residential area for the students. (2) the 8:30 am to 11:00 and 2:30 p.m. to 4:00 p.m in the two periods, removal of judgment 1 area, the rest were focused on 2, 3, 4, 5, 6, thus we can judge area 2, 3, 4, 5, 6 should be the student to study, the classroom, the technology center or to the library. (3) but at 6:30 p.m to 9:00 p.m, except area 1, the rest of the people are concentrated in area 6, according to the arrangement of the school, area 6 can be determined for the library. (4), however in area 2, 3, 4, 5,

the area 5 was significantly less than 2, 3, 4, due to the technology center is to do the experiment, is far more than the experiment and theory, so area 5 should be technical center, 2, 3, 4 areas should be 1,2,3teaching buildings.(5) at 11:00 to 12:30 a.m and 4:30 to 6:30p.m. to, everyone is focused on the area 7 and area 8, people in area 7 are far more than area 8, so area 7 maybe a dining hall, are 8 is snacks square.

Area 9 and area 10 people does not conform to certain rules, but on Tuesday afternoon we found people in area 9 and area 10 will be more than other working days, because the school arrangement is the school public holiday on Tuesday afternoon. So we speculated that 9 and 10 area should be student activities, basketball or football fields. On Sunday on Saturday, we found the most 3:00 PM to 6:00p.m students is area 11, we speculate that 11 area should be shopping district and leisure entertainment places. Another phenomenon is area 12 only after 4:30 in the afternoon on Sunday to Thursday night to somebody, but at the time of day of rest on Sunday on Saturday area 12 is only someone after 5:00 in the afternoon, because the school bathroom have corresponding is the open time, so we speculated that 12 area is in the bathroom.

Conclusion

The future of the world is not only oil drive, the data will also drive the world, science and technology and data fusion, along with the development of data, companies will follow the user's data, and is no longer a user on the one hand, choose and accept products enterprises. Digital campus is one of the important part of digital earth. Through the study of this lesson. In our view, we research the information can be in the school construction, to provide some reference.

1) now many universities in depth and digital campus information system is under construction, this is the reality the extension and expansion of the digital campus data space. At the same time, as an important part of the city, the campus is an urgent need to provide information and space or the location of the area, including improving teachers' utilization, rational utilization of limited resources on campus, students, administrators and school daily management and support, as well as the teaching and research work.

2) as the change of urban landscape. Business district, the district, and other external environment changing, school resources allocation should also change, such as school, school bus pickup points, etc., especially some old campus is the need to adjust.

3) Provide ideas and reference for the development of the corresponding software.

References

- [1] Yan-wei chai, zhao ying, Ma XiuJun, zhang yan "based on the behavior of the mobile positioning data acquisition and the geographical application research" in December 2010 Article number 1003-2363 (2010), 6-0001-07
- [2] Yunchen Jiang, Shanshan Wan Computer and Information Application (ICCIA), 2010 International Conference on the Issue Date: 3-5 Dec., 2010
- [3] Liu lin, ji-xian zhang, xin-ming tang, wan-wu li 2007 LBS system structure and key technology research article number 1009-2307 (2007), 05-0144-03
- [4] Zhang yongchuan. Based on the analysis of mobile phone positioning data users travel law, 2013
- [5] YuLe. Based on the research of location services of mobile phone positioning system, 2013
- [6] Wang Wei. cellular mobile station research of wireless location technology, 2006
- [7] Jian-qin zhang, QiuPeiYuan zhi-jie xu, ming-yi du. A travel itinerary identification method based on mobile phone positioning data. Journal of wuhan university of technology, 2013
- [8] ChaoFen. Smartphone location tracking system design and implementation, 2012

- [9] Wang lijun. Mobile phone positioning technology application in the traffic flow forecasting and research, 2010
- [10] Wu Yizheng, based on dynamic OD phone location information and travel survey acquisition method, 2014
- [11] Progress clustering Gong Xi, temporal trajectory, 2011
- [12] Chen Jia, Hu Bo, Zuo Xiao Qing, Yue Yang. Using a mobile phone positioning data mining user characteristics. Wuhan University, 2014
- [13] Ma yun, Hannover Fair speech, 2015
-