GPS Taxi Traces for Multi land Usage Classification of Extravagant

[1.] R. Saranya, [2.] R. Rajalakshmi, [3.] S. Karpagam, [4.] P. Ragunathan
[1.] M.E. (Embedded Systems)
[2, 4.] M.E. (Embedded System Technology)
[3.] AP/ECE, CMS College of Engineering

***Abstract*— The multi land use classification is done by collecting data of multi taxi from multi land. These data are collected for one year by fixing GPS in that taxi, which can be used to find the social function of the multicity. Detailed multi land use classification, which is difficult to obtain, is an integral part of urban planning. Currently, GPS traces of vehicles are becoming readily available. It conveys human mobility and activity information, which can be closely related to the multi land use of a region. First, we found that pick-up/set-down dynamics, Extracted from taxi traces data. Second, we found six features, which are designed to characterize the pick-up/set-down pattern of multicity. By using best combination of these six features we can improve the classification result. And also the latitude and longitudes data from GPS is widely used in order to get better classification result. The project discussed about "Multi land-Use Classification Using Taxi GPS Traces "from cities and that taxi. We discussed for three cities of taxi tracking for simulation in form graphs and this graph shows comparison among the city and shows which city is using taxi heavily for transportation. It can be done in order to avoid traffic jam.**

I. INTRODUCTION

Multi land use classification is an important aspect of urban planning. The taxi traces using global positioning system gives information about the mobility in that city. The concept of land use and land cover is used here. It means the area used and covered by GPS taxi. Early research on land and multi-land use classification attempted to recognize different ecological vegetation such as forests and wetland. Such land and multi land use classification has broad application in ecology, and relationship between urbanization and deforestation and farmland changes. Then classified urban land into built-up and non-built up lands to delineate urban region and model urban growth. More detailed classification of urban regions has usually focused on land-covers such as water, railway, and green lands. Most urban land-use classification research has used remote sensing data especially satellite images. The very first satellite used for monitoring the earth surface was launched by NASA in 1972 with resolution of 79m.The satellite resolution has increased gradually. The increase of satellite resolution promoted the development of new land use classification algorithms. And it led to better classification of land portion. The pixel based classification got worse result when the satellite based image resolution increased.

The multi land classification will be different for different city. Different classification of multi land use has been proposed with different objectives and application. The remote sensing techniques were used in order to get taxi traces data. But currently in this paper we are using large scale real world taxi trace data of multicity instead of single city. And also here we are collecting data from multi kind of taxi instead of single kind of taxi.

Here we first use clustering techniques in multicity to partition the city map into various zones. Second based on the passenger pick-up and set down dynamics pattern we can verify the social function of the multicity (multi land).then based on the social function of the test area and information such as traffic flow density.

We characterize the activeness of the area.

- Previous land use classification research is based on the physical properties of studied objects in remote sensing data. And the land use classification was done only for the single city (single land) and the passenger pick-up and set-down dynamics pattern are extracted only from that particular city. Currently the multi land use classification is based on the passenger pick-up/set-sown dynamics extracted from real world taxi trace data of different kind of taxi in multi land portion.
- The passenger pick-up/set-down dynamic pattern will be different for different land that is different multi city. For example Beach have more number of people visited in the evening than the morning. And scenic spots have more passengers in daytime than at night and have more passengers during holidays rather than weekdays.
- We designed six features extracted from the pick-up/set-down data of different time lengths. Each of the features was evaluated to recognize the region with social function of that region.
- We find that social function of regions is not steady all the time. And it may vary differently for different land portion. Here human flow and latitude, longitude data from global positioning system taxi trace gives activeness of all kind of regions in different multi land portion. The six features are given by

II. Related Works

This section briefly gives information about related works on taxi trace data and also for multi land use classification. The taxi trace data could be used for 1) Rebuilding a city road map. 2) Providing information about traffic condition, analyzing potential traffic hotspots. 3) We can identify routes for navigation because taxi trace data sometimes imply the experienced taxi driver knowledge.4) it helps people to avoid traffic jams.5) determining an abnormal trajectory since some taxi driver may be intentionally choosing a longer route to make more of a profit.

Mobility data may also contain information that can be used for developing intelligent transportation systems (ITS). The previous works on urban land use classification usually employ remote sensing techniques.

Previous works on urban land-use classification usually employ remote-sensing techniques. They differ from this paper in two ways. First few works considered fine grained classification of urban social function. Some of the previous research focused on land-cover classification of two kind of social functions, namely residential and non-residential .some works also defined roads like highway, railway, and so on. And some other previous research classified social function as residential, commercial and industrial and finally institution. None of the previous research gives information of land use and classification as this paper does. Since remote sensing data cannot depict enough information. Second, few works can handle the social activeness of all kind of region. Previous work depicted the activeness information by visual analysis of building density from remote sensing images. But this method of visual analysis has two limits 1) visual analysis of building density is not suitable for the building-scale regions, although it may be suitable for region with lot of builidings.2) it does not depict the activeness of human social activeness since building density is not closely related to human mobility. There are few approaches using mobility data for land and multi land use classification. The human mobility data can be obtained by using cell phone data, even the rural area of the city region can be covered, but we have to face serious problem in environment. Traffic level is increased. And agriculture land portion also reduced. For example we take a study area in china. Fuyang country is located in the south of Yangtze delta region & is one of the fastest growing industrial & an urban region in china.so the agriculture land loss is increased. It was apparent that the amount of agriculture land loss showed large fluctuations between 1999 and 2006. In 1999, 128 ha of agriculture land were converted to other uses. In 2000, that value changed to 133 ha, and then in 2001, it abruptly increased to 76 ha, and remained high in 2002 and 2003. It then abruptly decreased to 8.6 ha in 2004. And increased to 354 ha in 2005 and 217 ha in 2006. The agriculture land portions are converted to transportation, village residence, city and town and also for industry area.

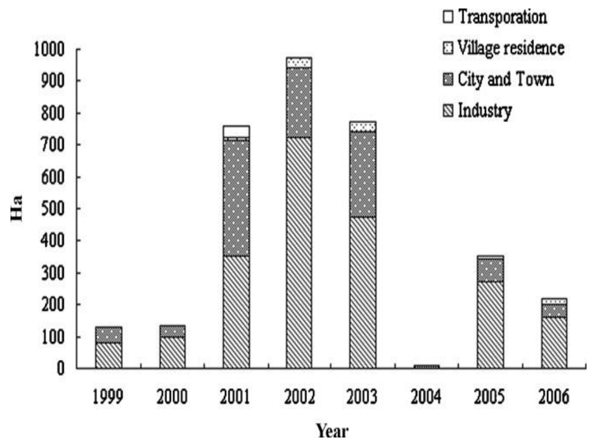

Figure 1: yearly agricultural land loss from 1999 to 2006.

Figure 1 shows that large gaps existed in the amounts of agricultural land converted to transportation, village residence, city and town, and industry.

Figure 2: Pattern and structure are important elements in image analysis beyond single pixel interpretation.

Origin-Destination (OD) flow gives information about human mobility & urban dynamics in city. LBSN, which is used for location based service. It also used in city planning and traffic engineering. It explored the relationship between Origin-Destination flows and social function of origins and destination to mine the semantics of OD flows. Cost is high and inefficient. Knowledge about OD is limited. Another research about finding traffic anomaly by using microscopic traffic variable. By using traffic variables we can able of achieve Transient, Traffic & incident originator with 100%detection rate, zero false alarm rate achieved. The loop detector used here has fixed location so it takes long time to find the traffic difference.

To achieve flexible and intelligent control of traffic and transportation systems, Wang developed an agent based networked traffic-management system. The agent-based control decomposes a sophisticated control algorithm into simple task-oriented agents that are distributed over a network. The ability of dynamically deploying and replacing control agents as needed allows the network to operate in a "control on demand" mode to adapt to various control scenarios. The system architecture employs a three-level hierarchical architecture. The highest level performs reasoning and planning of task sequences for control agents; the middle level dispatches and coordinates control agents; and the lowest level hosts and runs control agents. The control agents are represented by mobile agents that could migrate from remote traffic control centers to field traffic devices or from one field device to another. Traffic and transportation systems consist of many autonomous and intelligent entities, such as man-driven vehicles, signal lights, and variable signs, which are distributed over a large area and interact with each other to achieve certain transportation goals.

The agent architecture consists of five types of modules: *Sensors*, *Effectors*, *Communication*, *Motivation*, and *Cognition*. *Sensors* perceive the environment; *Effectors* take actions; the Communication module is responsible for communication with other agents; the *Motivation* module models the agent's long-term goals, roles,
and preference, whereas the *Cognition* module monitors and controls the agent's individual, communicative, and cooperative activities.

III. CHARACTERISTICS OF TAXI GLOBAL POSITIONING SATELLITE TRACE DATA SET

The taxi GPS Trace data set used in this paper of multi land usage classification comes from multi city in India. The taxi GPS traces were generated over a period of one year. During this period the GPS is installed in the taxi of multi city (land).

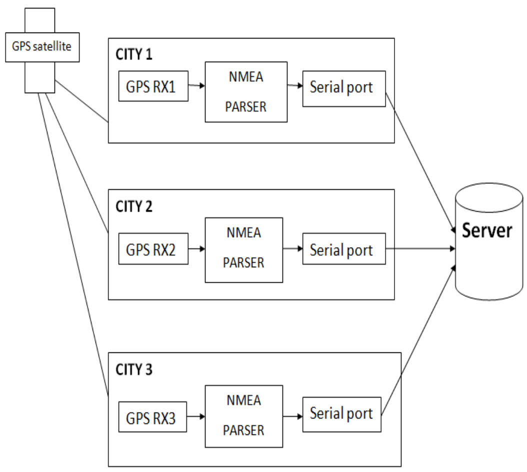

Figure 3: Functional diagram of Taxi Trace

The data set from GPS consists of the following information:

TAXI ID: Each taxi has unique ID.

TIME: The GPS data give time stamp "YYYY–MM–DD-HH:MM:SS"

GPS POSITION: It gives current longitude and latitude.

SPEED: Gives current Taxi Speed in km/h.

TAXI ORIENTATION: Gives the direction of the taxi.

METER STATE: It indicates whether the taxi meter is running.

Figure 4: pick-up numbers in a local area near the Hangzhou railway station. The color bar on the right side illustrates different colors for different pick-up numbers.

Set down event is extracted from the mass of taxi traces with the sampling and extracting pick-up/set-down events.

IV. REGION EXTRACTION: DBSCAN-BASED CLUSTERING

Regions with high pick-up density are informative in conveying human mobility cues, which is related to the social function of regions. This paper focuses on regions with high pick-up density. From the algorithmic view, these regions can be defined as convex hulls of clusters within the set of pick-up/set-down points. The challenge of extracting clusters comes from the disparity of clusters.

V. LAND-USE CLASSIFICATION

The pick-up/set-down dynamics is to describe the variation of pick-up/set-down number over time. First, we reveal the inherent relation between land-use classes and temporal patterns of pick-up/set-down dynamics. Second, to evaluate how useful the pick-up/set-down dynamics would be for land-use classification, six features are designed. Four typical classifiers are then employed to recognize the social function (land use) of regions. Moreover, we explore which features when combined could achieve the optimal classification performance. Moreover, we evaluate how much time of historical data the feature extraction requires to obtain reasonable classification accuracy. The six features are given by

- Daily pick-up features
- Daily set-down features
- Pick-up/set-down different features
- Pick-up/set-down ratio features
- Weekly pick-up features
- Weekly set-down features

VI. Region Dynamics and Social Activeness

A region's social function does not remain consistent all the time. It may change when buildings in the region are disused for reconstruction or change their utilization. For example, the land-use class of a region will change from station to residential district if the train station in the region is replaced by apartment buildings. We call such dynamic change of land use over time as region dynamics. The land-use classification result of a region over time can be used to depict the region dynamics. We employed the time length of 1 mo for feature extraction and classified land use of all the regions of each month. We found that the land use of many regions changed stably during the period of April 2009 to March 2010. Transition of social function is usually caused by unusual social events, such as traffic regulation, building reconstruction, and new center opening. For five transition examples, we further investigate the underlying social events that cause a change in land use.

VII. Conclusion and Discussions

In this paper, we explore and prove the potential use of taxi trace data in multi land-use classification. The data used in this paper are large-scale real-world taxi trace data of a big city with a population of 6 million. First, an improved clustering algorithm (iterative DBSCAN) is presented to extract regions, According to the characteristics of the data. Second, the land-use classification using the taxi trace data is proposed. Six kinds of Features are designed, and four classifiers are integrated into the evaluation. The performance of different features and classifiers is evaluated, and a best feature combination is also achieved. With the large-scale data, our approach can achieve an accuracy of 95% for land-use classification. Third, the dynamic transition of a region's land use can be detected and reveal corresponding social events.

Our work still has some limits for land-use classification. First, we cannot address regions that have few taxi passengers.

The taxi passenger flow is only a small part of the whole human flow and this result in some regions having fewer passengers. However, if the trace data of personal cars are available, our method can be easily applied to complementary trace data to handle more regions. Second, our work currently only addresses regions with pure land use. We do not consider regions with multiple land-use classes, which will be focus of future work.

References

[1] F.Sadowski and I. Sarno, "Forest classification accuracy as influenced by multispectral scanner spatial resolution, Environ.Res.Inst. Michigan, Ann Arbor, MI, Rep. 109600-71-F, 1976.

[2] R. Latty and R. Hoffer, "Computer-based classification accuracy due to data spatial resolution using a per-point vs per field classification techniques," in *Proc. Mach. Process. Remotely Sensed Data Symp.*, 1981,pp. 384–393.

[3] H. Geist and E. Lambin, "What drives tropical deforestation? A metaanalysis of proximate and underlying causes of deforestation based on subnational case study evidence," Land Use and Cover Change, Rep. Ser. No. 4, 2001.

[4] Y. Xie, Y. Mei, T. Guangjin, and X. Xuerong, "Socio-economic driving forces of arable land conversion: A case study of Wuxian City, China," *Global Environ. Change*, vol. 15, no. 3, pp. 238–252, Oct. 2005.

[5] S. Chen, S. Zeng, and C. Xie, "Remote sensing and GIS for urban growth analysis in China," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 5, pp. 593–598, May 2000.

[6] R. Gluch, "Urban growth detection using texture analysis on merged landsat TM and SPOT-P data," *Photogramm. Eng. Remote Sens.*, vol. 68, no. 12, pp. 1283–1288, Dec. 2002.

[7] D. Ward, S. Phinn, and A. Murray, "Monitoring growth in rapidly urbanizing areas using remotely sensed data," *Prof. Geogr.*, vol. 52, no. 3,

[8] C. Lo and Y. Xiaojun, "Drivers of land-use/land-cover changes and dynamic modeling for the Atlanta, Georgia metropolitan area," *Photogramm. Eng. Remote Sens.*, vol. 68, no. 10, pp. 1073–1082, Oct. 2002.

[9] W. Ji, J. Ma, R. Twibell, and K. Underhill, "Characterizing urban sprawl using multi-stage remote sensing images and landscape metrics," *Comput. Environ. Urban Syst.*, vol. 30, no. 6, pp. 861–879, Nov. 2006.