



2017 International Conference on Identification, Information and Knowledge in the Internet  
of Things

## A Method for Extracting Section Information of Highway Based on Massive GPS Data

Xinxin Zhang<sup>a</sup>, Guicheng Shen<sup>b,\*</sup>

<sup>a,b</sup>*School of information, Beijing Wuzi University ,Beijing 101149,China*

---

### Abstract

Information of highway is the basic element of intelligent transportation system. This study proposes a method for extracting section information of highway based on massive GPS data. The method has the advantages of low time complexity and high accuracy. The method is inspired by the algorithm of Density Based Spatial Clustering of Applications with Noise (DBSCAN). DBSCAN can discover cluster of arbitrary shape. The method calculated neighborhood density of GPS points by algorithm of zone division; therefore this method reduces the computation scale and the time complexity greatly. Taking into account the uneven distribution of GPS data between sections, the method of hierarchical division is adopted in clustering, and iterative clustering is computed from dense to sparse. Finally we find that our method can find the road information efficiently on massive GPS data.

Copyright © 2018 Elsevier Ltd. All rights reserved.

Selection and peer-review under responsibility of the scientific committee of the 2017 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2017).

*Keywords:* Extracting Road Section; GPS Data; DBSCAN; Algorithm of Zone Dividing; Iterative Clustering.

---

### 1. Introduction

There are many methods for extracting road information, such as manual mapping and analysis of remote sensing satellite image. The implementation of these methods usually cost high cost and long cycle. Today, the development of big data and GPS devices provides a new way for road information extraction. When the vehicle is running on the road, the GPS device will record the location, time and speed information. By analyzing these GPS data, we can obtain the road information.

---

\* Corresponding author. Tel.: +010-89534639; fax: +010-89534639.

*E-mail address:* [guichengshen@126.com](mailto:guichengshen@126.com)

Taxis as a convenient vehicle for people to travel, have the characteristics of the long travel time and wide range. The GPS data taxis generated is very suitable for highway information research. The GPS data records abundant spatiotemporal information, which can reflect the geometric characteristics of the highway. Thus, there are a lot of highway extraction methods based on GPS data. Edelkamp [1] extracted road information by the K-Means method. This method clusters the trajectory into the central line of the road formation path, but it is weak in noise immunity. Lee [2] propose a trajectory clustering method that divides and tracks trajectory data first. The method divides the whole trajectory and then analyzes the trajectories of the clusters. Kumar [3] extraction of road information by clustering GPS points with DBSCAN and improved clustering method. But for massive GPS data, using DBSCAN algorithm to extract the road is facing two major problems.

This study improves the DBSCAN algorithm of density clustering method. Different from other clustering methods, DBSCAN algorithm can identify clusters of arbitrary shape, and can filter out outliers and noise data. The algorithm is very suitable for road extraction.

## 2. Methods

- Algorithm of zone dividing: Zone division is to reduce the scale of calculation.
- Algorithm description of DBSCAN: DBSCAN can extracting section information of highway based on massive GPS data.
- Iterative clustering: This method can set the parameter flexibility, and improved the accuracy of the algorithm.

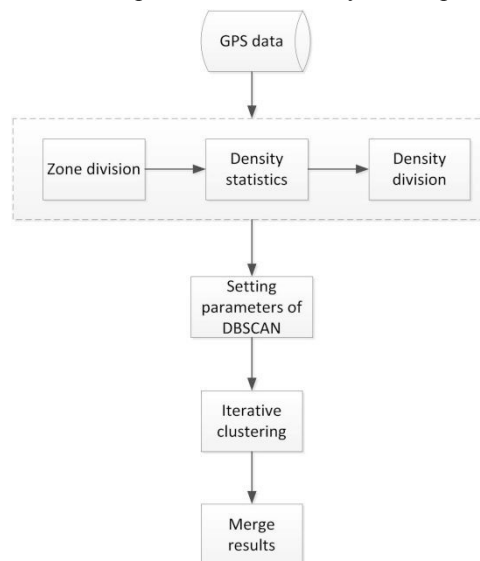


Fig.1. Flow chart of model

## 3. Data Preparation

The original data is 100 thousand taxi trajectories records, and the capacity is about 1G. A record of GPS points is generated with in 15s in the process of vehicle running. Data fields include trajectory ID, vehicle ID, service type, timestamp, track point latitude and longitude.

In this stage, we changed 100 thousand trajectories into 4770835 GPS points. This process mainly includes splitting the trajectory, cleaning noise data and converting data type.

Download English Version:

<https://daneshyari.com/en/article/6900201>

Download Persian Version:

<https://daneshyari.com/article/6900201>

[Daneshyari.com](https://daneshyari.com)