Anchor Points Seeking of Large Urban Crowd
Based on the Mobile Billing Data

Wenhao Huang, Zhengbin Dong, Nan Zhao, Hao Tian, Guojie Song*, Guanhua Chen, Yun Jiang, and Kunqing Xie

Key Laboratory of Machine Perception, Ministry of Education, Peking University, Beijing, 100871, China
{rubio8741,znzn007,tianhao86}@gmail.com,
{dongzhengbin,gjsong,kunqing}@cis.pku.edu.cn,
{ghchen,jiangyun}@pku.edu.cn

Abstract. In everyday life, people spend most of their time in some routine places such as the living places (origin) and working places (destination). We define these locations as anchor points. The anchor point information is important to the city planning, transportation management and optimization. Traditional methods of anchor points seeking mainly based on the data obtained from the sample survey or link volumes. The defects of these methods such as low sample rate and high cost make it difficult for us to study on the large crowd in the city. In recent years, with the rapid development of wireless communication, mobile phones have becoming more and more popular. In this paper, we proposed a novel approach to obtain the anchor points of the large urban crowd based on the mobile billing data. In addition, we took advantage of the spatial and temporal patterns of people’s behavior in the anchor points to improve the simple algorithm.

Keywords: mobile phone, anchor point, OD information, spatial-temporal pattern, large crowd.

1 Introduction

There are some places which we stay most of our time in daily movement and everyday living such as living places (Origin) and working places (Destination). We commute between workplace and home. These places are referred as the anchor points. Obviously everyone has his or her own anchor points. Those anchor points are of significant value to effective traffic network management, public transportation planning, city rational planning and many related social aspects [1]. For example, the travel between living place and working place which is also called commuting is the most basic and common travel for everyone [2]. It has crucial impact on the optimization of urban transportation especially the morning and evening rush [3].

* Corresponding author.
Traditionally, we obtain the anchor points information by the following two methods: survey method and modeling method [4]. However, though these two methods own their own merits, they are far from satisfaction. In recent years, modern wireless communication especially the mobile phone services have become more and more popular and widespread all over the world. The coverage of mobile phone is larger than 90% to the whole population in European Union [6]. Almost 8 billion people in China own a cell phone. This means a surprisingly full coverage of mobile phone using globally. For the purpose of billing, the billing data recorded by the mobile phone operator often include the location information such as the base station coordinates. This provides us a great opportunity to use these data to seek the anchor points of the large crowd.

In this paper, we proposed a simple algorithm to obtain the anchor points of the large crowd based on the mobile phone billing data. In order to overcome the irregular correlation between the calling patterns and special-temporal patterns of anchor points, we take full advantage of the spatial-temporal stability of the anchor points to improve accuracy and reliability of the algorithm.

2 Related Work

2.1 Anchor Points Estimation Techniques

Researches on the anchor point estimation could be divided into three categories according to the methodology [7].

![Fig. 1. OD estimation techniques](image)

These techniques all have their own advantages and disadvantages. Sample survey conduct traffic survey by mail or household interview. The data collected by the survey are very detailed and accurate and could be used to many other related works. However, due to the cost and time, the sample rate is quite low. Meanwhile, the survey area is often limited in some near districts or blocks. Furthermore, the error will be increased greatly in the conversion process from the surveyed samples to large crowd.

The estimation by link volumes has large amount of data [8]. But it is only part of the crowd and is constrained coverage area. At the same time, error caused by the assignment model restricted its widespread use.

The wireless communication method has emerged in recent years. It is becoming a hot topic in transportation especially the intelligent transportation systems [9]. Though the data is not very detailed because it is designed for billing not for our purpose, it is