Journal of Transport Geography 66 (2018) 309–320

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations

Kyoungok Kim

Information Technology Management Programme, International Fusion School, Seoul National University of Science & Technology (SeoulTech), 232 Gongreungno, Nowon-gu, Seoul 139-743, Republic of Korea

ARTICLE INFO

Keywords:
Public bicycle system
Negative binomial regression analysis
Clustering analysis

ABSTRACT

Public bicycle systems are widely spread across many cities worldwide. 'Tashu', a public bicycle sharing system in Daejeon, was installed in 2009 and it is one of the well-established public bike-sharing systems in South Korea. Previous studies in the literature found that in general, bicycling is affected by weather conditions and temporal characteristics. However, the degrees of impacts or the signs of effects may be different depending on the stations. Therefore, this study investigated the different effects of weather conditions and temporal characteristics according to the characteristics of the stations at the station level analysis in addition to the system level analysis. For the cost-effective station level analysis, clustering analysis was utilized to find out the groups of the stations with the similar properties. Moreover, temperature humidity index (THI) and the indicator variable of heatwaves were introduced to consider the interaction between temperature and humidity and measure the influence of high temperature, which has been rarely considered. In the system level analysis, the results showed that the selected factors have the different influence over the different time periods within a day. Especially, scorching heat and non-working days differently affect the demand for public bikes by hours. Also, it was observed that high temperature over 30 °C reduces the bicycle usage, which revealed the necessity of taking into account not only severe colds but also heatwaves in the prediction of the demand. By clustering analysis, the stations were partitioned into the three clusters. One cluster shows the strong peak in the morning while two others have peaks in the evening. The effects of weather conditions and non-working days on the demand for public bicycles were different depending on the clusters, which seemed to be related to the main purposes of bike usage in the clusters.

1. Introduction

A public bicycle system, also called as a bike-sharing system is a service in which bicycles are made available for shared use. It allows people to borrow a bike from one of the stations belonging to the system and return it at another station of the sharing system. The public bicycle system began in Amsterdam in 1965 (Shaheen et al., 2010), but the expansion of the system to the world was due to the development of IT technology in the 2000s. The development of technology has made it easier and more convenient for people to rent a bike, and mobile IT devices have allowed people to find a nearby station where they can rent a bike. In 2014, public bike-sharing systems were distributed over 50 countries on five continents and included 712 cities (Shaheen et al., 2014; Ahillen et al., 2016).

Such a bike-sharing system provides an affordable alternative to motorized public transport and private vehicles for short-distance trips in an urban area and aims to reduce noise and air pollution. In addition to this, it was reported that transit use increases in cities with bike-sharing due to improved connectivity to other means of transit (DeMaio, 2009). However, the bike-sharing system has been criticized as less convenient than a privately owned bicycle and as relatively inefficient and high cost in relation to its effectiveness. Therefore, to ensure the success of a bike-sharing system, it is necessary to select appropriate areas for the stations and predict the demand according to different situations.

Current bike-sharing systems are supported by information systems that automatically collect records related to renting and returning of bikes, and this information provides useful data for the bike-sharing system. Utilizing this information not only predicts the long-term demand at a station, it can also determine the relationship between the local characteristics and external factors and the demand. Therefore, analyzing the rich data of the detailed trip records of the system can be used to successfully manage the bike-sharing system. In addition, this information can be used to understand the human mobility patterns that are influenced by bicycle use in urban areas, which may show characteristics that differ from the mobility patterns by other modes of transit.

E-mail address: Kyoungok.kim@seoultech.ac.kr.

https://doi.org/10.1016/j.jtrangeo.2018.01.001
Received 11 July 2017; Received in revised form 7 December 2017; Accepted 2 January 2018

In recent years, research utilizing such data has been actively conducted. Several studies investigated the spatio-temporal variations in demand across a system to predict the demand at the system level or station level. Those studies provided ways for optimizing the location of stations and the inventory management or found the factors that affected the demand for public bicycles. The different works used different approaches or models and different explanatory variables to estimate the demand for public bicycles. In Kaltenbrunner et al. (2010), prediction of the total number of available bicycles or free slots in the stations was conducted using an Auto-Regressive Moving Average (ARMA) model for the community bicycle program, Bicing in Barcelona. Since ARMA is a tool to understand a stationary stochastic process in the analysis of time series, this work only utilized the data which consists of the total amount of bicycles available at the stations measured at the same time interval in modeling for prediction. Unlike this work, many other studies usually used additional explanatory variables in prediction. García-Palomares et al. (2012) built the models to predict the potential demand for sharing bicycle as the first step for optimizing the location of stations of the bike-sharing program in Madrid. The spatial distribution of the demand was estimated using the population and employment. Corcoran et al. (2014) studied the effect of weather and calendar events on the total number of trips by all stations of CityCycle in Brisbane using a Poisson regression model. However, this work did not provide the statistical models in sub-system level or station level. Instead, the spatio-temporal dynamics of public bicycle trips in sub-system level were visualized by flow-comaps. Rixey (2013) investigated the effect of demographic and built environment on average monthly bicycle usage of the three cities in USA using linear regression. As similar to Rixey (2013), Faghih-Imani et al. (2014) investigated the effect of land use and urban form on arrival and departure rates at the stations of BIXI in Montreal by regression analysis, but it used a linear mixed model instead of a simple linear regression model to consider station specific effects.

The several kinds of literature used various information, including the geographical characteristics of stations, such as land use, built environment, and accessibility, or external conditions such as weather and calendar events, to explain the public bike ridership in detail. However, it is not easy to find out research that considered the possibility that the selected factors are differently related to the pattern of public bicycle usage. In many cases, the models were trained using the local characteristics of stations as station specific explanatory variables and the coefficients of the factors were estimated as they have the same influence on the demand regardless of stations. Therefore, these models could not reflect the interaction between the characteristics of stations and other factors.

To address this issue, this study used clustering analysis to group the stations into homogeneous subsets, because building a separate model for each station is costly and not efficient although the separate models can measure the different effects of explanatory variables depending on stations. On the other hand, learning a separate model for each cluster is much efficient and it is possible to identify the relationships between the common characteristics within clusters that help to understand human mobility by bicycles and other explanatory variables on the demand for public bicycles.

In this study, I investigated the effects of external factors such as weather and non-working days on public bike ridership as similar to the literature, using the bicycle usage data obtained from the bike-sharing system called 'Tashu' operated in Daejeon, South Korea. This study estimated the separate negative binomial models for the detected groups of the stations by clustering analysis as well, in addition to the system level models. The temporal patterns of people renting public bicycles were used to measure the similarity between the stations with reference to other literature which found that the temporal variations in human mobility are generally associated with the geographical characteristics of regions (Handy, 1996; Dieleman et al., 2002; Pan et al., 2013; Zhang et al., 2014). The bicycle flows between the detected clusters are utilized in order to characterize the clusters and derive plausible interpretations of the estimated models. Additionally, I introduced new variables, temperature humidity index (THI) and the indicator variable whether the highest temperate exceeds a certain point or not. The first variable is introduced to deal with the interaction between temperature and humidity. The second variable is generated to reflect the phenomenon that outdoor activities are limited in the hot summer, while other studies have shown that the demand for public bicycles increases as temperature increases.

The rest of the paper is organized as follows. Section 2 introduces the study area, Daejeon and explains the datasets in detail. Afterwards, the methodology used for the analysis in different levels was described in Section 3. Then, Section 4 provides the results of the system and station level analysis. Finally, the conclusions of the article is presented in Section 5.

2. Case study region and datasets

This section first describes Daejeon, the study area, and then introduces the datasets used in this study. The aim of the study was to reveal the relations between weather and calendar events on the bicycle ridership at the system and station levels. I conducted an exploratory data analysis and then completed the appropriate preprocessing steps needed to build the models necessary for the aim.

2.1. Geographical area covered by the study

Daejeon is the fifth largest city in South Korea and located in the central region of South Korea. Due to the characteristics of this location, Daejeon serves as a hub of transportation. The area of Daejeon is 539.97 km^2 and the population was approximately 1.5 million in 2015.[1]

Daejeon maintains policy efforts to promote the low carbon growth. As part of that effort, Daejeon has been pursuing several policies to expand bicycle use in order to save energy, eliminate traffic congestion, improve the environment and promote citizen's health. The policies to expand bicycle use include a public bike-sharing system and provisions to create and maintain bicycle paths and bicycle facilities as well as to create incentives for transferring public transportation to bicycles. As of 2014, the total length of bicycle paths in Daejeon was 107.6 km and the length of the shared roads by pedestrians and bicycle riders was 627.8 km.[2] Moreover, a public bike-sharing system was introduced in 2009 as a pilot project and after the pilot project, the system has been extended to other areas besides the pilot project area. In addition to the bicycle infrastructure, Daejeon has a nationwide high rate of bicycle use. As a result of the survey on the use of bicycles by the Korea Transport Institute in 2016, Daejeon had the highest proportion of people riding a bike more than once a week (36%) in South Korea.[3]

2.2. Datasets

2.2.1. Tashu trip data

Tashu is the self-service public bike share system in Daejeon.[4] It was launched in 2009. As of May 2017, the system had 226 stations and 2955 racks throughout Daejeon. Bike rental is possible for both members and non-members. When registering as a member on the website, users can pay a flat fee for a week, a month, or a year, or purchase a one-day pass when they want to take a bicycle. For non-members, the one-day pass can be purchased after a confirmation process using a

[1] Based on the population census conducted in 2015 by the commissioner of the Statistics Korea.
[2] Based on the statistics by Daejeon Metropolitan City (https://www.daejeon.go.kr/sta/index.do).
[3] https://www.koti.re.kr.
[4] https://www.tashu.or.kr/.