Review of Literature on analyses of Bike Sharing Systems

Abstract

We review several papers and online websites and applications on bike sharing systems (BSS).

Introduction

Several works have been devoted to analyses of bike sharing systems (BSS), and have lead to several applications to streamline the use of these systems. We review these here, in several sections. There are mainly two categories of papers: the exploratory type which study the data and the insights it provides about the system usage, and preditive type which try to predict demand. One problem for which several solutions have been offered in literature is that of station balancing.

Data analysis

O'Brien, Cheshire, and Batty paper (O'Brien, Cheshire, and Batty 2014).

Mainly interesting for: defines and studies certain measures of BSS usage.

The load factor is the proportion of docking points in each docking station that currently have a bicycle available to hire. The basic statistic for each docking station is expected to be the number of bicycles and the number of free spaces in each docking station.

They work with data obtained from APIs (normally Google Maps API), a collection of pin-style markers, representing the docking stations, whith the capacity and load factor appearing as attached statistics. System operators seem to provide these APIs. In some cases third party APIs, often run by volunteers.

structure of various bike-shares

1. aggregate characteristics that provide simple (non spatial) measures

- max number of stations,
- max number of bicycles in the stations
- biggest change between the daily max and min number of bicycles

2. spatial characteristics that look at the placing of docking stations within each system

- latitude of the centroid of the system
- system's area of influence
- the observed mean distance between stations
- the Z-score (which describes whether the system's footprint is statistically clustered, random or dispersed)
- the compactness ratio (system's shape)

3. temporal characteristics

- the load factor
- a normalized measure of the redistribution needed to even out the load
- how these measures vary on an intra-day and weekday/weekend basis.

Maximum concurrent usage for day

$$U_{dmax} = (B_{dmax} - B_{dmin})/B_{dmax}$$

, where B_{dmax} is the maximum bicycles available in day, and B_{dmin} the minimum.

Load factor should be related to the uni-directionality of flow.

If charges are time-based, non-availability of a docking station at the end of a trip is worse than non-availability of a bike at the start. Load factors are around 50% for many BSSs.

Fig. 3: No. of bikes available against time shows a strong time dependent pattern. Ripe for Fourier analysis.

Spatial Characteristics

A BSS can be large in area, have a high/low density of docking stations, subnetwork density may vary between core and periphery.

Clustered systems – Z-score Z-score computed for the distribution of distances between all the stations normalized by the distance between each point and the central mean point of the system. Is this distribution random, or exhibits signs of clustering or dispersal.

Clustered systems have a large negative Z-score, dispersed systems have a large positive Z-score. Randomly distributed systems have a near zero Z-score.

System classification based on temporal characteristics

- Commuters: home to office
- Utility Users: shopping, errands
- Leisure Users: weekends for fun and exercise
- Tourists Users: go to beach or explore city

Cities can be clustered by their hourly usage patterns.

The Station balancing problem

Several papers studying the balance of stations, and optimal re-balancing operations.

Demand prediction models

Resource allocation models

References

O'Brien, Oliver, James Cheshire, and Michael Batty. 2014. "Mining Bicycle Sharing Data for Generating Insights into Sustainable Transport Systems." *Journal of Transport Geography* 34 (0): 262–73. doi:http://dx.doi.org/10.1016/j.jtrangeo.2013.06.007.