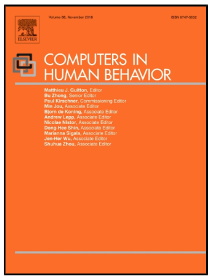
Accepted Manuscript



Customer Attractiveness Evaluation and Classification of Urban Commercial Centers by Crowd Intelligence

Haixia Mao, Xiaopeng Fan, Jinping Guan, Yeh-Cheng Chen, Haoran Su, Wenzhong Shi, Yubin Zhao, Yang Wang, Chengzhong Xu

PII:	S0747-5632(18)30394-7
DOI:	10.1016/j.chb.2018.08.019
Reference:	CHB 5653
To appear in:	*Computers in Human Behavior*
Received Date:	13 April 2018
Accepted Date:	10 August 2018

Please cite this article as: Haixia Mao, Xiaopeng Fan, Jinping Guan, Yeh-Cheng Chen, Haoran Su, Wenzhong Shi, Yubin Zhao, Yang Wang, Chengzhong Xu, Customer Attractiveness Evaluation and Classification of Urban Commercial Centers by Crowd Intelligence, *Computers in Human Behavior* (2018), doi: 10.1016/j.chb.2018.08.019

Customer Attractiveness Evaluation and Classification of Urban Commercial Centers by Crowd Intelligence

Haixia Mao
School of Automotive and Transportation Engineering, Shenzhen Polytechnic, 518055 China
7098, Liuxian Road, Nanshan District, Shenzhen, Guangdong, P. R. China, 518055
Tel: (86) 18576646026, Email: maohaixia@szpt.edu.cn

Xiaopeng Fan, Corresponding Author
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen, P.R.China
Tel: (86)86392329; Email: xp.fan@siat.ac.cn

Jinping Guan
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences;
1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen, P.R.China
Tel: (86) 82950655; Email: melon_ping@163.com

Yeh-Cheng Chen
Department of Computer Science, University of California, Davis, CA, USA.
Email: ycch@ucdavis.edu

Haoran Su
Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong
Email: shr1992@163.com

Wenzhong Shi
Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong
Email: john.wz.shi@polyu.edu.hk

Yubin Zhao
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
Email: zhaoyb@siat.ac.cn

Yang Wang
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
Email: yang.wang1@siat.ac.cn

Chengzhong Xu
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

Email: cz.xu@siat.ac.cn

Customer Attractiveness Evaluation and Classification of Urban Commercial Centers by Crowd Intelligence

***Abstract*—Evaluation on urban commercial centers' attractiveness not only benefits business strategy making and location choice, but also helps traffic management and urban planning. Traditionally, it is studied using questionnaires and field research, which are labor-intensive and time-consuming. To overcome these problems when evaluating the urban commercial centers' attractiveness, massive data analytics with datasets from taxi traffic, population, area, and road networks are adopted in this paper. Taking fifteen commercial centers at Shenzhen as a case study, a Cyber-Physical-Social System is built up to deal with these massive data for statistical analysis. An "Attractiveness Degree Model" is proposed to describe the degree to which customers desire to visit a commercial center. Then attractiveness thematic maps are drawn. Results show that YiTianJiaRiGuangChang has the highest attractiveness degree even though it has a small size and low commercial value. The attractiveness degree rankings are corroborated by annual customer satisfaction survey from Shenzhen Retail Business Association. Attractiveness thematic maps show that about 50-65% visits by taxis are within 5 km range. These results can be applied to support market analysis, urban planning, traffic management, and related areas.**

***Keywords*—Urban commercial centers; attractiveness evaluation; attractiveness degree model; Cyber-Physical-Social System; massive data analytics**

I. INTRODUCTION

Identifying urban commercial centers' attractiveness is very important. It not only benefits business strategy making and location choice, but also helps traffic management and urban planning. In previous research, it was studied using the central place theory [1] [2], gravity assumptions [3] [4], discrete choice models [5] [6], spatial interaction models (such as entropy-maximizing models [7]), radiation models [8][9], statistical analysis (such as regression) etc. Data is traditionally collected through field research [10], which has a high cost of manpower and time.

With the rapid developments of modern cities and sensing devices, big cities result in massive data. The attractiveness of urban commercial centers can now be evaluated based on massive data analytics rather than traditional questionnaires. Massive data analytics helps to overcome field research's labor shortage and time consuming.

In massive data analytics, GPS trajectory data plays a very important role. Many studies were based on these float trajectories in urban computing for different applications [11-17]. Trajectory data from a wearable device was used to determine the user's task and predict his future movements [11]. In that study, the proposed theory clustered all GPS data automatically and incorporated them into a Markov model. Trajectory data from taxi with passenger loaded or unloaded information was used to evaluate taxi movements, analyze idle taxis, support taxi management, and provide real-time traffic information etc. [12-15]. Studies mostly use the data in transportation field.

In other fields, GPS trajectory data can represent human movements and further help to define a region's function in urban computing. Using GPS trajectories and Point Of Interests (POIs), Zheng et al. presented a framework to discover the functions of different regions in a city [16]. Using taxi GPS trajectory data, Qi et al. found that the amount of taxi passengers who got on or off the taxi in a specific region could describe the social activity dynamics in that region [17].

Another interesting field is to evaluate the attractiveness of urban commercial centers using taxi GPS trajectory data. So far, only some studies have looked into this issue. Kawasaki and Axhausen used GPS data to generate choice set for grocery shopping location choice model [18]. Yue et al. presented a method to evaluate the attractiveness of shopping center using the generated trips and the catchment areas [19] [20], based on the taxi trajectories in Wuhan, China. However, there are two points that should be reconsidered in the last study. In that study, there were over 63,000 trajectories from around 12,000 taxis. This means that one taxi had approximately only five trajectories per day, which was much lower than the overall Wuhan taxi report [21]. This shows that the dataset was not completed. So the accuracy of the results must be negatively influenced. Huff's model was calibrated using taxi GPS trajectories to delineate shopping center trade areas [3] [20]. New findings argued that the traffic flux between two locations should take population into consideration [8] [9]. Therefore, using taxi GPS trajectory data, an attractiveness degree analysis of urban commercial centers is needed, in which population should be included as a factor.

In this paper, all the taxi data gathered from Shenzhen are used as the full-sampling dataset, and the population as a factor is introduced when studying commercial centers' attractiveness.

Customers travel to urban commercial centers in all different kinds of ways, including private vehicle, bus, metro, walking, bike, and taxi etc. With taxi trajectory data, this study focuses on the by-taxi visitors. It evaluates urban commercial centers' attractiveness of customers traveling by taxi based on massive data analytics. In future research, transit card data and parking data will be used to evaluate other modes like bus, metro, and private car etc. Then the overall attractiveness of different urban commercial centers will be more accurate.

This paper uses crowd intelligence by massive data to evaluate the attractiveness of commercial centers at Shenzhen, one of central cities in China with a population of approximately 12 millions. 15 commercial centers are selected in NanShan District at Shenzhen as our case area, which are the de-facto commercial centers recognized by the customers from public credibility organizations, such as the Dianping website, which is the most popular website for POIs in China. Based on the GPS trajectories and transactions from 15,000 taxis, a series of statistical analysis is carried out, including taxi visits (pick-ups and drop-offs for weekday or weekend and for different periods of time) and taxi cost (taxi fare, travel time and travel distance). And the insights into the classification of commercial centers based on massive data statistical analysis is presented.

The main contributions can be summarized as follows:

- The concept of commercial center is precisely defined geographically and sociologically;
- The "Attractiveness Degree" model is proposed to evaluate the attractiveness of commercial

centers;

- A Cyber-Physical-Social System is introduced to deal with massive data from three spaces for crowd intelligence;
- Massive data statistical analysis and thematic maps shed new insights into the commercial centers' attractiveness;
- The commercial centers are classified into representative types based on their evolutions and massive data statistical analysis.

The rest of the paper is organized as follows. Section 2 discusses the related works. In Section 3, urban commercial center is defined, and the attractiveness degree model is presented. In Section 4, massive data processing is explained and the proposed Cyber-Physical-Social System is introduced. In Section 5, massive data statistical analysis is conducted, including taxi visits and taxi cost. And the insights into commercial centers' classification based on the massive data statistical analysis are presented. In Section 6, the attractiveness degree model's results are discussed and attractiveness thematic maps' results are presented. Finally, this paper is concluded in Section 7.

II. RELATED WORKS

Researchers from all over the world have worked on the topic of the attractiveness of urban areas for many years. The seminal work of the gravitational attractiveness model is proposed by Reilly in 1931[4]. Christaller et al. present the theory of central places in 1933[1]. Based on the two works above, new methods for evaluating the attractiveness of urban areas are presented from various research fields, such as urban planning, transportation, marketing, business, immigration and traveling etc. These works introduce new impact factors, such as local population, distances, surveys, trajectories, and time etc.

Traditional methods on evaluating urban area attractiveness include surveys, geographical features and population distributions. For example, Huang et al. present that the attractiveness of POIs mainly depend on some static factors, such as the area of commercial center, the distance between customer residence and commercial center etc., but sometimes dynamic factors like meal time also have impacts on the attractiveness of restaurants [36].

Recently, Dolega et al. present a flexible model with a composite index of attractiveness, considering the impact of the interdependencies between different retail centers [40]. However, this method involves a series of simplifications when expanding the assumptions used to model catchments for a store chain or retail/service category to retail agglomerations. Ortegón-Cortázar et al. consider that the design of ecological spaces and environments has the potential of becoming a field of interest for shopping centers with a potential effect on visiting and shopping intentions [41]. However, the results depend on costly collecting the samples comprised of 449 consumers from 25 different shopping centers in Bogota. Mittal et al. discuss the attractiveness of shopping center in Indian and consider there are four dimensions, including merchandising, variety & selection, milieu & facilities, and convenience [42]. However, this work did not present a clear model to describe the relations among these four factors.

A new method by investigating the trajectories of mobile objects like vehicles, mobile phones, or persons to evaluate the attractiveness of commercial areas is presented. Giannotti et al. propose a new algorithm to evaluate the attractiveness of an urban district by the number of mobile objects traversing it [37]. Giannotti et al. also improve this work by building up a grid for segmentations of customer interests and combining the neighboring cells with similar features [38]. In the framework of dealing with trajectories data, Wei et al. present a framework based on pattern-aware trajectory search [39]. The proposed algorithm is based on trajectory density to extract the districts traversed by a specific number of trajectories. Then the attractiveness of these districts can be deduced by random walk.

In the field of analysis on taxi trajectories, Yue et al. present a method to evaluate the attractiveness of shopping center by combining and clustering similar taxi trajectories in terms of drop-offs and pick-ups, and propose to take time dependency among clusters as the standards of clustering [19]. Next, Yue et al. take into consideration some new metrics, such as total rented area, the number of shopping centers and parking slots etc., to evaluate the attractiveness of shopping centers [20]. All the data are from taxi trajectories in Wuhan, China.

In the field of analysis on pedestrian trajectories, Furtado et al. leverage the time of pedestrian activities in a district, i.e. Stop-at, Pass-by and Pass-in, to evaluate the attractiveness of the district [39].

TABLE I EXISTING WORKS FOR REVIEW

Existing Works	**Trajectories of Individuals**	**Scale of Analysis**	**Categories of Places**	**Trajectory Episodes**	**Support of Episodes Parameters**
[4] The gravitational attractiveness Model (1931)	N/A	Single	Predefined	N/A	N/A
[1] The central place Theory (1933)	N/A	Single	Predefined	N/A	N/A
[36] The spatial temporal POIs' attractiveness model (2010)	Any	Single	Predefined POIs	Stops and Passing	No
[40] The multi-staged spatial interaction model (2016)	N/A	Multiple	Predefined	N/A	N/A
[41] The structural equation model (2017)	N/A	Multiple	Predefined	N/A	N/A
[42] The mall-intercept survey (2016)	N/A	Multiple	Predefined	N/A	N/A
[37] The sequential pattern mining paradigm (2007)	Any	Single	Created	Passing	No
[38] M-Atlas (2011)	Any	Single	Created	Passing	No
[39] PATS (2010)	Any	Single	Created	Origin/ Destination	No
[19] OD analysis (2011)	Taxi	Single	Shopping	Pick-up/	No

			mall only	Drop-off	
[20] OD analysis (2012)	Taxi	Single	Shopping mall only	Pick-up /Drop-off	No
[43] M-Attract (2012)	Any	Multiple	Predefined /Created	Stops and Passing	Yes
Our work	Any	Multiple	Predefined /Created	Pick-up/ Drop-off and Cost	Yes

In general, we review the existing works in Table I. Existing methods for evaluating the attractiveness of commercial areas can be described in terms of 5 dimensions, including the trajectories of individuals, the scale of analysis, the categories of places, the trajectory episodes, and the support of episodes parameters as well. Especially, how to divides a trajectory into episodes is very important for attractiveness evaluation. It is necessary to divide a trajectory with a series of tuples (Longitude, Latitude, Timestamp) into a sequence of episodes according to a given predicate, and provide support for semantic analysis. Compared with existing methods, our proposed A-degree model for the first time considers the datasets, including transportation traffic, population, road network, GIS, real estate prices, and website etc. from the cyber space, the physical space, and the social space as well. The relationships among these factors are accurately described by the A-degree model. The ranking is confirmed by the customer satisfaction survey from Shenzhen Retail Business Association.

III. Definitions, Models, and Datasets

A. Definitions

In this paper, the concept of "urban commercial center" is defined as one or more existing nearby commercial buildings with its definite borders and vertices. It is a commercial area, within which customers can accomplish a series of activities by walking, such as shopping, dining, and entertainment etc. It can include urban complexes with offices or residence upstairs.

A commercial center (also called commercial area) is denoted as $CA[V_1, V_2, ..., V_n]$, in which each of the vertex is in the form of [Latitude, Longitude]. 15 commercial centers are selected as the study case in NanShan District, Shenzhen (Figure 1 and Table II). These commercial centers are the de-facto commercial centers that are recognized by the customers, and the most popular centers in Nanshan, one of the biggest districts at Shenzhen. They are distributed and also cover almost the whole area of this district. The criteria of selection are based on the de-facto standard. These centers are selected by the ranking and comments from customers at the Dianping website (http://www.dianping.com/), which is the best website for POIs in China.

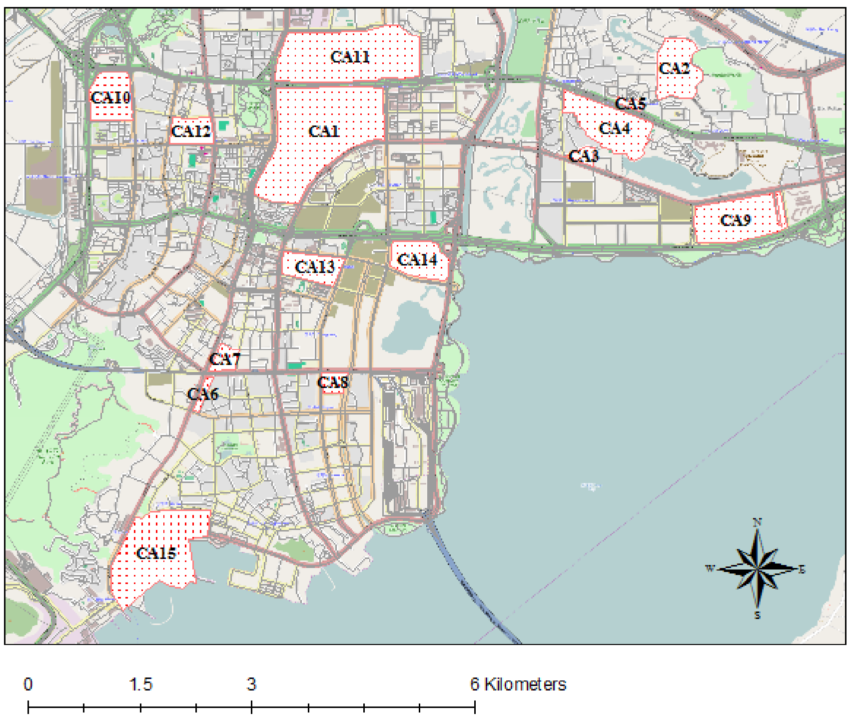


Figure 1. 15 commercial centers in NanShan District, Shenzhen

TABLE II LIST OF COMMERCIAL CENTERS

No.	Name	Area(m^2)	Population
CA1	ShenzhenUniversity	1542228	24757

CA2	HuanLeGu	371526.4	4511
CA3	JingJiBaiNa	51945.4	601
CA4	WindowOfTheWorld	430733	3707
CA5	YiTianJiaRiGuangChang	30688.07	318
CA6	HuaYuanCheng	35389.56	2225
CA7	XinHeChunTian	93418.95	1119
CA8	BaoNengTaiGuCheng	77776.7	931
CA9	HuanLeHaiAn	575203.1	800
CA_{10}	ZhenYeXingHai	315088.6	5731
CA_{11}	KeJiYuan	1158117	8392
CA_{12}	TianHong&HuanLeSong	1748.91	2814
CA_{13}	NanShanCentral	253781.6	3376
CA_{14}	ShenZhenWanTiYu	300663.3	300
CA_{15}	HaiShangShiJiei	1037727	19831

Among these commercial centers, CA1 can be taken as an example. There are 7 vertexes in CA1: (x1, y1), (x2, y2), (x3, y3), (x4, y4), (x5, y5), (x6, y6), (x7, y7). Each commercial center can be a concave or convex polygon in the map. The edges of a polygon are usually the borders of a commercial center. And the border is composed of the surrounding roads with their buffers. The buffer size is defined as 70 meters.

B. Model

By using customers' visits by taxi, a new model with the name attractiveness degree model is presented, i.e. A-degree in formula (1), to evaluate the urban commercial centers' attractiveness degree of the customers who are traveling by taxi. It is an index for a commercial center which is attracting customers who travel by taxi. The A-degree of a commercial center describes the popularity of a commercial center for customers to visit by taxi.

$$A - degree_i = \frac{\dfrac{V_i * r_1}{cost(i) * \left((P_i - c_1)^{\alpha} + c_2\right) * r_2}}{\sum_{i=1}^{n} \dfrac{V_i * r_1}{cost(i) * \left((P_i - c_1)^{\alpha} + c_2\right) * r_2}} \quad (1)$$

All of the key terms in the new model in this paper are listed as follows.

TABLE III TERMS AND DEFINITIONS IN THE MODEL

CA_i	The i[th] commercial area
$A\text{-}degree_i$	CA_i attractiveness degree of customers traveling by taxi
V_i	Total visits by taxi of CA_i
r_1	Average number of passengers in a visit (trip) per taxi
P_i	The population of CA_i
c_1	Average population of the fifteen commercial areas
c_2	Customers' visits radix in Nanshan District, Shenzhen
r_2	Taxi mode share of the customers
α	The magnification coefficient from population to visits
$cost(i)$	Taxi visit cost of CA_i(travel time, taxi fare and distance)
k_1	Taxi travel time coefficient in Principal Component Analysis
k_2	Taxi travel distance coefficient in Principal Component Analysis
k_3	Taxi fare coefficient in Principal Component Analysis

There are three major factors which have important impacts on the attractiveness, including the customer visits to the commercial center, the costs for the visitors and the population base within the commercial center. It is different from the previous works in this filed. The A-degree model considers the travel behaviors of customers by the visits in both spatial and temporal dimensions. In the A-degree model, the attractiveness of a commercial center is proportional to the visits, but is inversely proportional to the cost of trips to a commercial center. The population base is also considered in this model. More or less population will have negative impacts on the attractiveness. Excessive population base will keep customers from a commercial center due to the overcrowding problem and the safety issue, while smaller population base will make a location lack of popularity.

The rationality of the A-degree model is from an aggregate perspective. There are three key factors in the model to evaluate the commercial centers' attractiveness of customers who travel by taxi.

① Number of customers who visit the commercial center by taxi is at the numerator. Total visits by taxi to a commercial center(V_i) multiplied by the average number of passengers in a visit (trip) per taxi (r_1) is the number of customers who visit the commercial center by taxi. It reflects directly the attractiveness of a commercial center for customers traveling by taxi.

②The cost function of visiting a commercial center by taxi is c*ost(i)*. It is inversely proportional to the A-degree as shown in formula (1). It is estimated through Principal Component Analysis (PCA), with a function of taxi fare, taxi travel time, and taxi travel distance to a certain commercial center. Taxi fare is a function of travel time and travel distance, but not only limited to them. The reason is that taxi fare is also related to the travel schedule. For example, the taxi fare from A to B will not be the same if the trip happens at 8:00am and at1:00pm in a day. Traffic condition also has an influence on the fare. Thus, taxi fare is related to the travel time and travel distance, but also varied with time. These three factors in the function are all calculated from our taxi data.

All the taxi Origin-Destinations (ODs) are in the form of (travel time, travel distance, fare) as the samples, and Principal Component Analysis is used to estimate the values of (k_1, k_2, k_3) . Results show that it is

(0.0018, 0.3628, 0.9318). The detailed function is shown in formula (2).

$$cost(i) = k_1 * \text{ taxi travel time} + k_2 * \text{taxi travel distance}$$

$$+ k_3 * \text{taxi fare} \quad (2)$$

③ Population base and its visits to a commercial center by taxi is included in formula (1).

In the Huff's model, the square footage and the travel time are the key factors to attract a customer to a shopping center [3]. However, in the cost-based radiation model, population becomes more and more important [8]. The average flux from one location to another mainly depends on the populations of two locations, the intersectional population between the two locations, and the distance between them. These factors are all included in the formula (1).

In this model, both the historical factor of a commercial center and the motivational factor of the visits outside of a commercial center are taken into consideration. The impact of the existing population base on the attractiveness of a commercial center is considered, which reflects each commercial center's historical factor, i.e, population grows and decreases over time. Over and under population both have a negative impact on the attractiveness [22] [23], which is the motivation factor of the visits outside of a commercial center. Over-population keeps customers from visiting a commercial center due to the overcrowding and unsafe issue, while under-population makes a location lack of popularity.

Parameter r_1 and r_2 can be evaluated using monthly statistic reports at the website of Shenzhen Transport Commission of Shenzhen Municipality (http://www.sztb.gov.cn/xxgk/tjxx/). From monthly reports, the parameter r_1 can be calculated by finding the average number of passengers in a visit trip per taxi, through dividing total taxi person-time by total taxi transactions in that month. The parameter r_2 is obtained by finding the proportion of taxi passenger among bus, taxi, and subway for public transportation. Their values are 2.5 and 11% from two years reports.

For the magnification coefficient α, an empirical value, this parameter is designed to control the influence of the residents in a commercial area on its attractiveness. In the A-degree model, "customers" (the same meaning to "consumers" in this paper) are different from the "residents". Our population base is from the WorldPop project and is to describe the behaviors of the residents who may also be the customers and live within a commercial center (This situation in China may be much different from that in Europe or US). Our objective is to investigate taxi data to describe the behaviors of customers, who may not be the residents within a commercial center. It is noted that the range of [1.0, 1.5] are much more suitable and the value of 1.1 is used in this paper.

C. Data

Our dataset is composed of five subsets: taxi GPS trajectory data, taxi transaction data, road network map, population map, and real estate price map. Taxi GPS trajectory data and transaction data are obtained from the embedded devices in each taxi.

① Taxi GPS Trajectory Data

It is from all the taxis at Shenzhen for 7 continuous days, including workdays and weekend. The data includes approximately 200,000,000 records for over 15,000 taxis. The total size is 48.5GB in the format of text. The frequency of GPS sampling is 20-60 seconds.

TABLE IV TAXI GPS TRAJECTORY DATA FORMAT

Name	**Type**	**Example**	**Memo**
Taxi_ID	String	粤 BXXXXX	
Longitude	Double	114.1X4634	
Latitude	Double	22.5X5097	
Timestamp	String	2014-05-20 00:00:16	
Device_ID	Integer	0000000	
Velocity	Double	43	km/h
Orientation	Integer	280	Heading direction
Position_Status	Boolean	1	Positioned

Passenger_ Loading_Status	Boolean	1	Loaded
Taxi_Color	String	Red	Red/Green/Yellow

②Taxi Transaction Data

All the taxis' transactions are included for the corresponding 7 days, which come from different taxi companies. Data format is as follows.

TABLE V TAXI TRANSACTION DATA FORMAT

Name	**Type**	**Example**	**Memo**
Taxi_ID	String	粤BXXXXX	
Start_Time	String	2014-05-19 23:39:40	
End_Time	String	2014-05-19 23:59:17	
Unit_Price	Integer	312	#/100
Mileage	Integer	15256	#/1000
Timesheet	String	00:00:00	Hour/Min./Sec.
Fare	Integer	5620	#/100
Deadhead Mileage	Integer	2898	#/1000
Consumer Card ID	String	1	
Original Residual	Integer	0	#/100
Remaining Residual	Integer	0	#/100
Frequency of Outage	Integer	0	
Duration of Outage	String	00hour00min00sec	
OverspeedingDistance	Integer	0	#/1000
Overspeeding Frequency	Integer	0	

Taxi License ID	String	00100000	

③Three Maps: Road Network Map, Population Map, and Real Estate Price Map

These three map datasets are included to evaluate our models. The road network map is obtained from the OpenStreetMap website [24], an open source data under the Open Database License. The map has pre-made shapefiles, roads, trails, cafés, railways, stations, etc.

The population map is achieved from the WorldPop project [25]. This project united the continent-oriented AfriPop, AsiaPop, and AmeriPop projects. It was initiated in 2013. It provides detailed and free-access population data and maps for Central and South America, Africa, and Asia.

The real estate price map is from the Urban Planning, Land and Resource Commission of Shenzhen Municipality [26].

IV. Massive data Processing: A Cyber-Physical-Social System

This section focuses on the framework that deals with the multi-dimensional data. Five datasets are included in this study to evaluate the urban commercial centers' attractiveness of customers, who travel by taxi. These datasets are from different dimensions, including areas, populations, road networks, real estate prices, taxi trajectories, and taxi transactions.

This study is based on a Cyber-Physical-Social System (CPSS). In the dataset, the areas and road networks are two geographical attributes of urban commercial centers. And they are converted to GIS data in the cyber space. The taxi GPS trajectories and transactions are human activities in the physical world. The populations and real estate prices are the two social attributes of commercial centers. These three sides form the Cyber-Physical-Social System. Such a system is built up to deal with these massive data in order to make analysis for applications, like the attractiveness of a commercial area, landscape visibility analysis [30], night bus route planning [31], on-bus Wifi passenger behavior analysis [32], residents characteristics analysis [33] [34] [35]. The Cyber-Physical-Social System can deal with a series of datasets and provide a unified platform to collect, clean, pre-process, manage, and analyze massive data.

The procedure of the massive data evaluation in this study is shown in Figure 2.

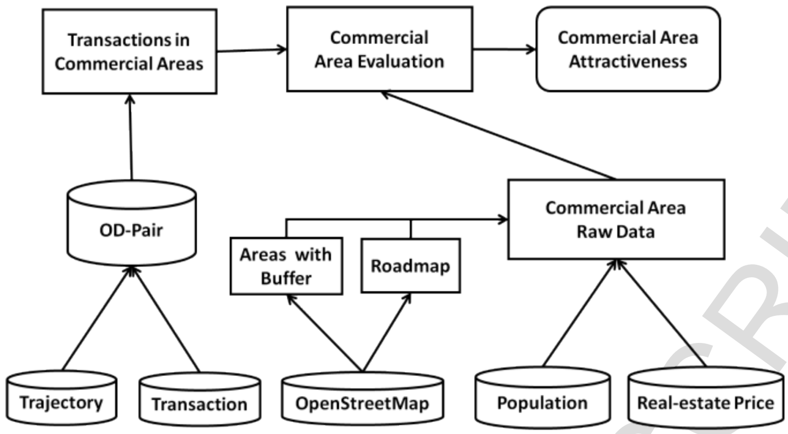


Figure 2. Massive data evaluation process (Five datasets are at the bottom.)

A. Data Pre-processing

Since errors or defects exist in the raw data of GPS trajectories and transactions, data pre-processing is firstly conducted. This ensures the quality of data itself and the further evaluation. Data pre-processing steps are as follows (Figure 3):

Step 1: Taxi Data Cleaning. Taxi data format mistakes are reformatted and some typical errors are removed from the two taxi datasets: taxi GPS trajectory data and taxi transaction data.

In the trajectory data, one typical error is trajectory missing. This can result from GPS signal loss, signal fading by tree or building covering, and transmission failure etc. Most of the trajectory losses are caused by transmission failures and may be reported in the following days. So one day complete taxi data is collected from its ten neighboring days for integrity. Another typical error is location drift. It is corrected by map matching.

In the transaction data, there are two typical errors: duplicate transactions and un-unified decimal points. Decimal points are not unified because of the different record formats of different taxi companies. To solve these two issues, only one copy for each transaction is kept and all the decimal points are unified. The unified transaction data format is shown in Table V.

Step 2: Data Extraction with Commercial Area's Definition. A commercial center is defined by its spatial area of surrounding roads and is delineated with a buffer of 70 to 100 meters. Within the area, there are buildings, Point-Of-Interests, and roads etc. All the vertexes are obtained and the area of a commercial center is clearly defined in ArcGIS. With the clear definition of each commercial center, its area, population, and real estate price are all extracted. Corresponding area is calculated in the form of square meters.

Population base data is obtained from the WorldPop project in which it is evaluated using remote sensing images [25]. Population base of each area is calculated by integrating all the points in the area and summing them up. Real estate price of a commercial center is evaluated by the average real-estate price within the range of the commercial center. Transaction records are found on the website of Urban Planning, Land and Resource Commission of Shenzhen Municipality [26].

Step 3: Origin-Destination (OD) pairing with Commercial Area's Definition. All the trips (OD pairs) to or from a commercial center (a commercial area) are extracted from the taxi trajectory and transaction data. The origins and destinations of each transaction are recovered, and the origins or destinations that are in the commercial center among all the taxi trajectories are found. Spark [27], one of the massive data analysis tools, is used to deal with these massive data.

B. Statistical Analysis

The whole procedure of statistical analysis is shown in Figure 3.

Pick-up and Drop-off Visit Analysis: All the visits can be classified into pick-ups and drop-offs. Pick-ups mean the trip's origin is within the range of the commercial center. Drop-offs are the ones that have the destinations within the range of the commercial center. Total visits of each commercial center are the sum of all the pick-ups and drop-offs.

Periods Analysis: For different periods of time of a day (0:00-6:00, 6:00-10:00, 10:00-18:00, 18:00-21:00, 21:00-24:00), pick-up and drop-off visits are analyzed.

Cost Calculation: each taxi visit's travel time, fare and distance for the cost function (Formula (2)) are calculated. And average taxi travel time, fare, and distance for each commercial center are analyzed.

Based on these results, commercial centers are classified into different categories.

C. Model Estimation

With the results above, the A-degrees for all the commercial centers are calculated. The ranking of the A-degrees is used to evaluate the urban commercial centers' attractiveness of the customers who travel by taxi.

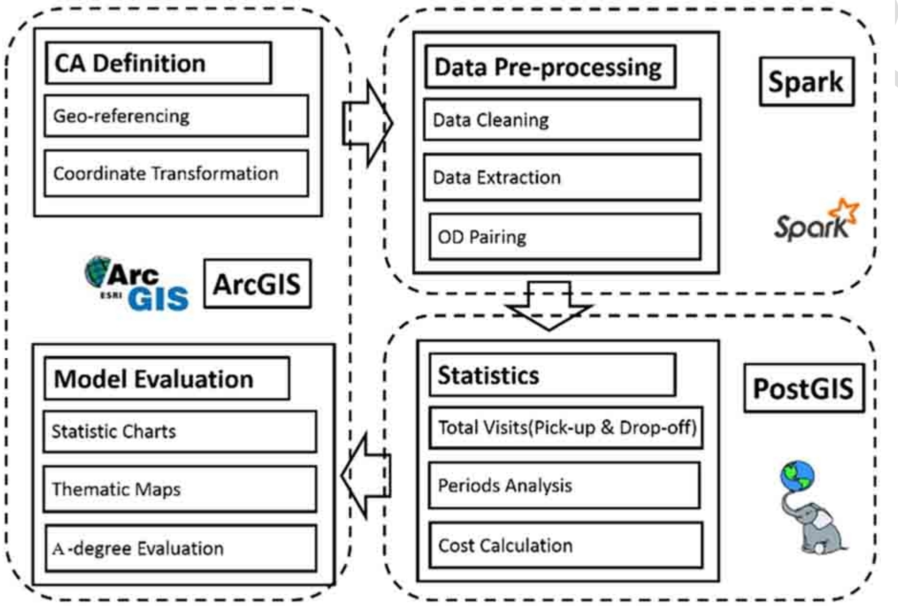


Figure 3. Data flow processing chart

Attractiveness thematic maps are drawn for the highest and lowest attractiveness degree commercial centers.

In this study, the Cyber-Physical-Social System has two subsystems. One is the GIS-based system. It is used to deal with the map data and other small data. The other one is the Spark-based system. It is used to process the massive taxi data.

The GIS-based system is based on the calculation of PostGIS and ArcGIS, in which there is an Object-Relational Database Management System with the support of spatial functions. Both geometry and geography data are supported in the spatial database. They can be treated in the plane coordinate system and the spherical coordinate system. The disadvantage is that the calculation is in the stand-alone mode. It has a high accuracy but a low efficiency.

With the data of the cyber-physical-social system (especially the traffic data) increasing, stand-alone mode can no longer efficiently support the data processing. Parallel and distributed computing models are introduced in order to deal with massive data. Spark is adopted as our major computing platform, which is one of the most influential platforms for massive data analytics. It is developed by the AMP Lab at UC Berkeley [27]. Both batch and streaming data can be iteratively in-memory processed in Spark.

Our experimental settings of the two subsystems are as follows. The GIS-based system is running on a personal computer with 2.5GHz 4-cores Central Processing Unit (CPU), 4G Random Access Memory (RAM), and 500G Hard Disk (HD). The Spark-based system is running on a cluster consisting of two high performance servers. Each node is with an Intel(R) Xeon(R) 2.4GHz 16-cores CPU, 8GB RAM, and 8T HD. 10 Virtual Machines (VMs) for our analytics are built up. Each VM is with 2.4GHz 4-cores CPU, 16G RAM, and 1T HD. They connects with each other by 1000M Ethernet and all the nodes run Community Enterprise Operating System 6.5 x86\64 Linux System.

V. Statistical Analysis

In this section, all the results of the statistical analysis are discussed, including visits (pick-ups and drop-offs) and cost (taxi fare, travel time and travel distance). And the insights on the classification of commercial centers are presented.

A. Taxi Visits (Pick-ups & Drop-offs) of Each Commercial Center

Taxi visits going into or out of a commercial center are one of the key factors to evaluate attractiveness. All the taxi ODs from taxi GPS trajectory and transaction data are collected.

All the related ODs of a commercial area are calculated. Total taxi visits (pick-ups and drop-offs) of each commercial center are shown in Figure 4. They are shown with the average of the weekdays and weekends respectively.

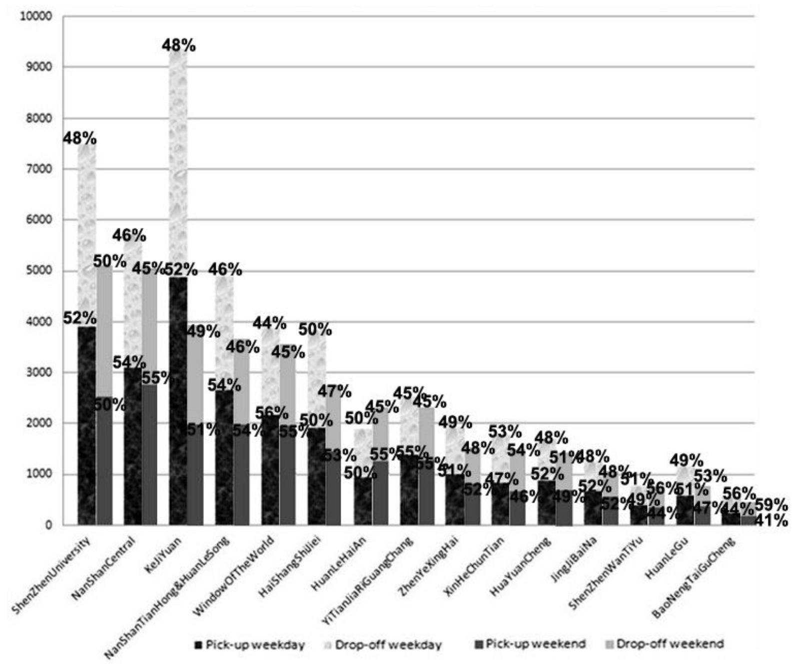


Figure 4. Total taxi visits (pick-ups and drop-offs) of each commercial center with the average of weekdays and weekends respectively

In Fig. 4, KeJiYuan is with the most taxi visits on an average weekday, while ShenzhenUniversity has the most taxi visits on an average weekend day. For most commercial centers, an average weekday has more visits than weekend. KeJiYuan has the greatest difference between weekday and weekend taxi visits. This can be explained by "a large portion of the visits to KeJiYuan on a workday are to the science and industry park for working or business purpose" [28].

Most of the commercial centers tend to have more taxi pick-ups than drop-offs. This is because comparing to taxi drivers having no control in drop-offs, it is much easier to find passengers to pick up in a commercial center and a taxi driver is inclined to find the next passage in these populous areas.

B. Periods Analysis of Taxi Visits (Pick-ups & Drop-offs) in Each Commercial Center

Whole day taxi visits are studied in the above section. In this section, the whole day is divided into five periods: the morning period (6~10am), noon and afternoon period (10am~6pm), evening period (6pm~9pm), night period (9pm~0am), and early morning period (0~6am). Taxi visits (pick-ups and drop-offs) are investigated in details by dividing the whole day into five periods.

During different periods of time of a day, various commercial centers have different characteristics. Some may mainly consist of offices. Therefore, most of the taxi visits may be in working hours. Some may have night club and karaoke, so that most of the visits are at the night.

Commercial centers have different functions, such as shopping, recreation, working, and habitation etc. Human activities reflect various major functions in various commercial centers. Different commercial centers have different kind of attractiveness for various customers and visitors.

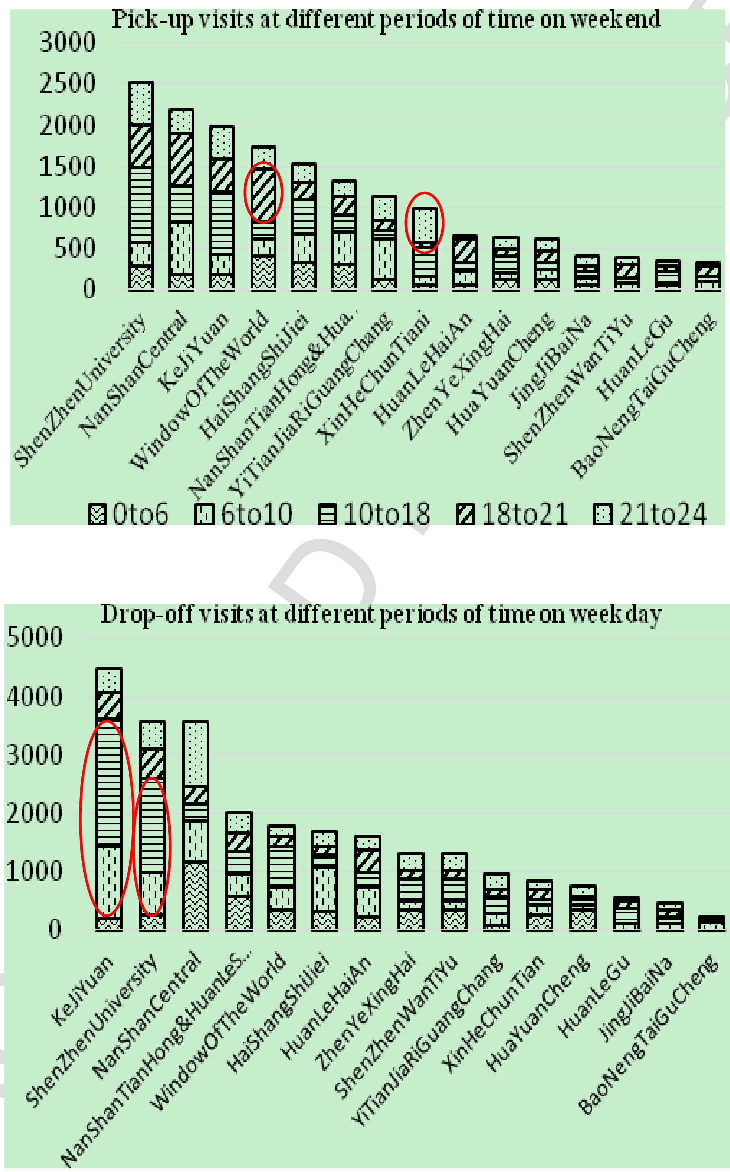


Figure 5. Rankings on taxi pick-ups on weekend and drop-offs onweekday during five different periods of time

Taxi pick-ups and drop-offs during different periods of time on weekday or weekend are all studied. Figure 5 only shows taxi pick-ups on weekend and drop-offs on a weekday.

Fig. 5 upper part shows that on weekend WindowOfTheWorld has a major portion of the pick-ups happening during evening period (6pm~9pm). As a popular theme park at Shenzhen, most of the visitors leave this center in this evening period.

Fig. 5 upper part also shows that on weekend in XinHeChunTian, pick-ups of night period (9pm~0am) are much more than other periods. This shows that XinHeChunTian has a major recreational function in the evening.

Fig. 5 lower part shows that on weekday, in KeJiYuan and ShenzhenUniversity, taxi drop-offs mostly happen during work hour (6am-6pm). This shows that KeJiYuan, as a commercial center with many office buildings, has a major function of business and working. And Shenzhen University has a major function of education, so most of the drop-offs happen during work hour, when classes take place.

In conclusion, our statistics on traffic data show that the trips to or from commercial centers should inosculate the trends of visiting demands' differentiation.

C. Travel Cost Evaluation of Commercial Centers

The cost for customers to visit a commercial center is evaluated in this subsection, including travel fare, travel time, and travel distance of taxis.

For all of 15 commercial centers, radar chart is presented to describe average travel fare, travel time and travel distance of taxi pick-ups and drop-offs on weekday (Fig. 6). Since weekend has no significant difference, it is not shown. Fig.6 shows that taxi cost to HaiShangSheJie is the highest among all the commercial centers.

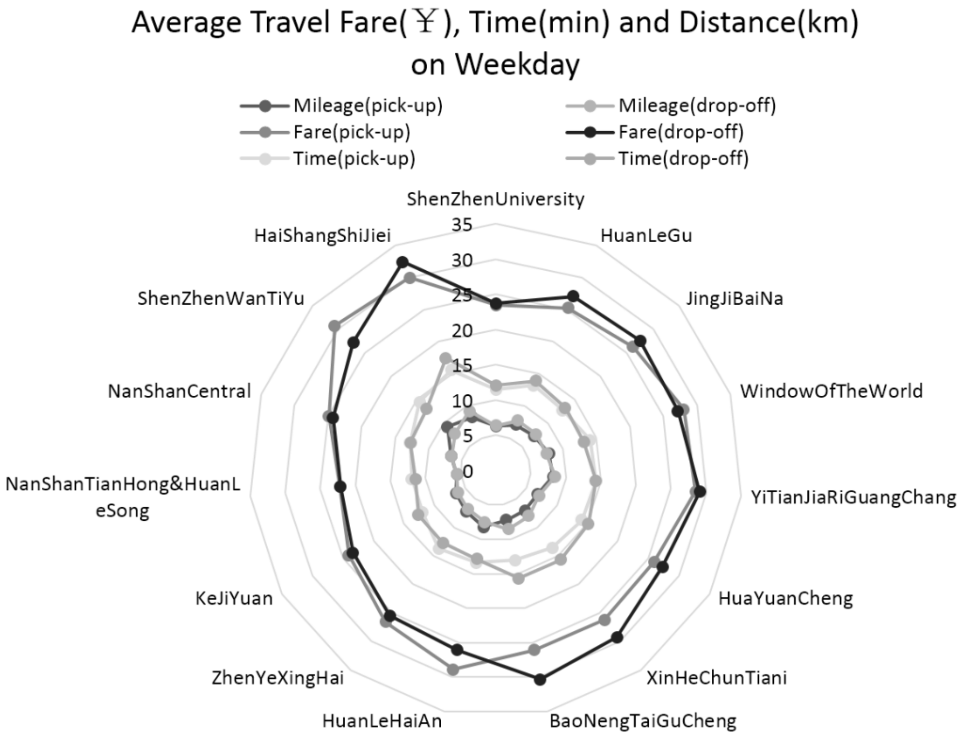


Figure 6. Average travel fare, time, and distance of commercial centers on weekday

D. Classification of Commercial Centers

In this subsection, 15 commercial centers are classified into 4 categories as follows:

- **Long-Term-Stable Type**. This kind of commercial center has long history of more than 10 years. Consumers know its popular name and normally it is with excellent public reputation. NanShanCentral, ShenZhenUniversity, KeJiYuan, WindowOfTheWorld, ZhenYeXingHai, HuaYuanCheng, and NanShanTianHong&HuanLeSong belong to this type. This type of commercial centers is normally with a stable residential population and a small migration population, such as a commercial center near a known university.
- **Sporadic-Burst Type.** This kind of commercial center has many public facilities, such as a famous stadium. When there is a public activity (such as a concert or tournament), there will be a burst of customers to congregate. But there are fewer people in the area most of the time. ShenZhenWanTiYu belongs to this category. This type of commercial centers normally is open for citizens to see a vocal concert or watch a sport match.

- **Day-Night-Contrast Type.** Between day and night, this kind of commercial center has obvious difference on visits or activities. HuanLeHaiAn, HaiShangShiJie and XinHeChunTian belong to this type. For example, in Fig. 5 upper part, XinHeChunTian's pick-ups of night period (9pm~0am) are much more than other periods on weekend. This shows that XinHeChunTian has a major recreational function in the evening. These commercial centers have diversiform activities at night. This type of commercial centers normally is not inhabitable and there will few visits after the facilities within them are closed.
- **Emerging-Comprehensive Type.** Commercial centers of this type have relatively new shopping malls, restaurants, and other facilities. YiTianJiaRiGuangChang, JingJiBaiNa, HuanLeGu, and BaoNengTaiGuCheng belong to this type. This type of commercial centers is emerging and promising in China with densely populated settlements. Normally, customers in this type can complete a series of activities with different objectives, such as shopping, dining, entertainment etc. This is the key point for the definition of commercial centers, presented in Section III.

VI. ATTRACTIVENESS DEGREE EVALUATION

The attractiveness degree (A-Degree) model is our core model to evaluate the attractiveness of commercial centers. In the third section of this paper above, a detailed definition for this model is presented. In this section, all the A-degrees of all 15 commercial centers are calculated and analyzed. The results for weekday and weekend are shown in Figure 7 and Figure 8.

The ranking shows that YiTianJiaRiGuangChang is with the highest attractiveness degree among all the commercial centers on both weekday and weekend. The most essential reason is that its location is the best among all the commercial centers. This commercial center is near WindowOfTheWorld and it lies within the Overseas Chinese Town. These two neighbors bring continuous passenger flows. YiTianJiaRiGuangChang exploits its land and provides high quality service.

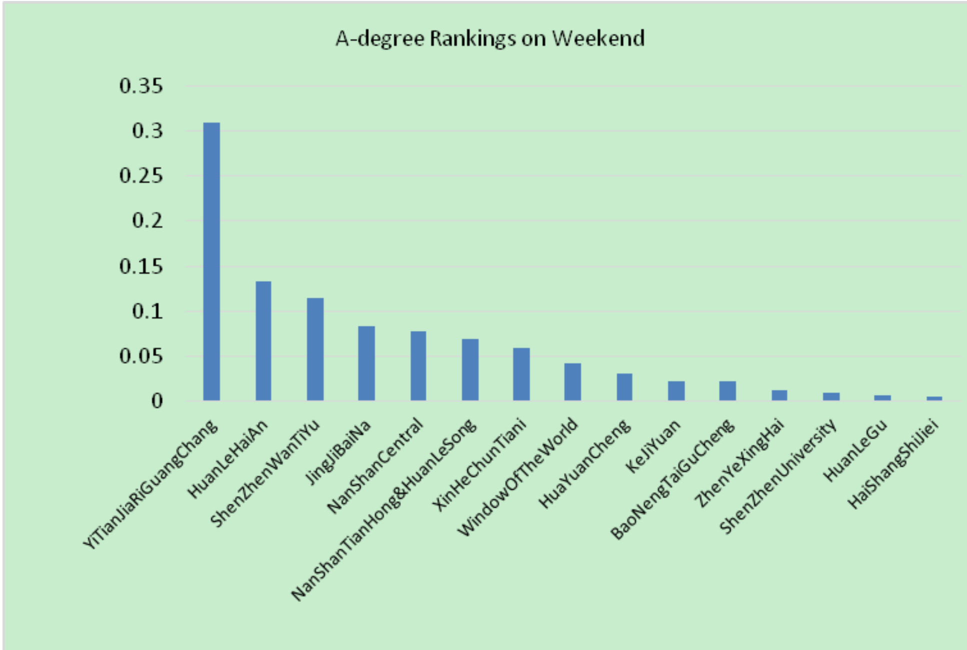


Figure 7. A-degree rankings of 15 CAs at weekend

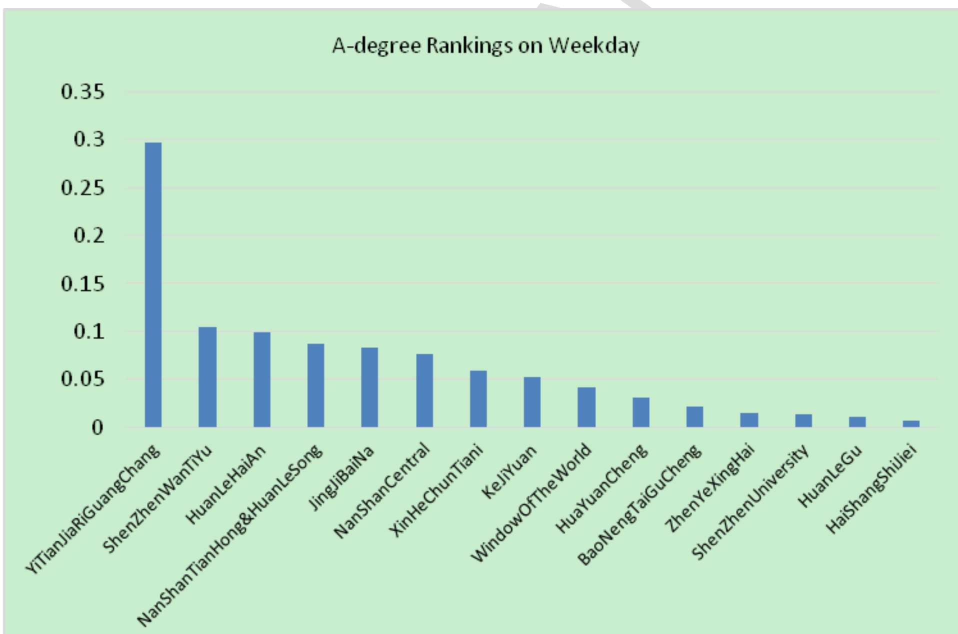
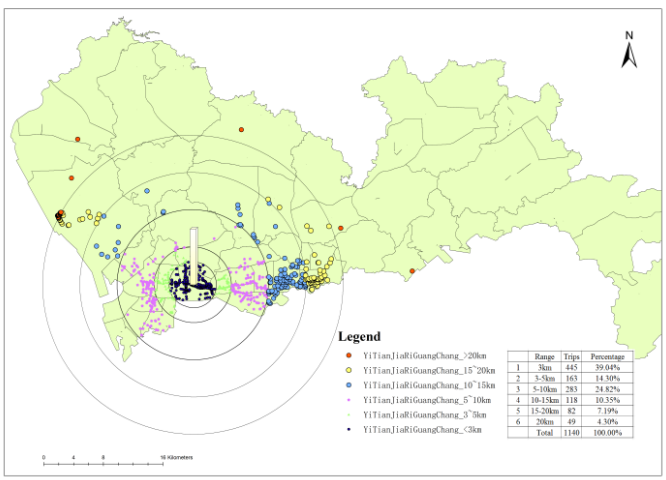
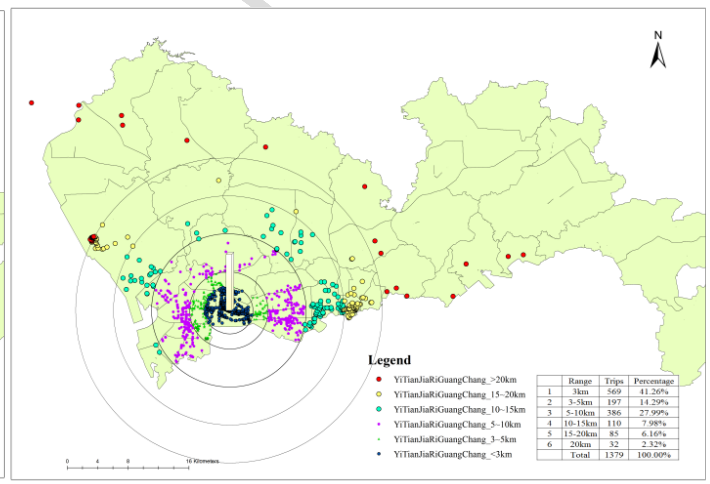


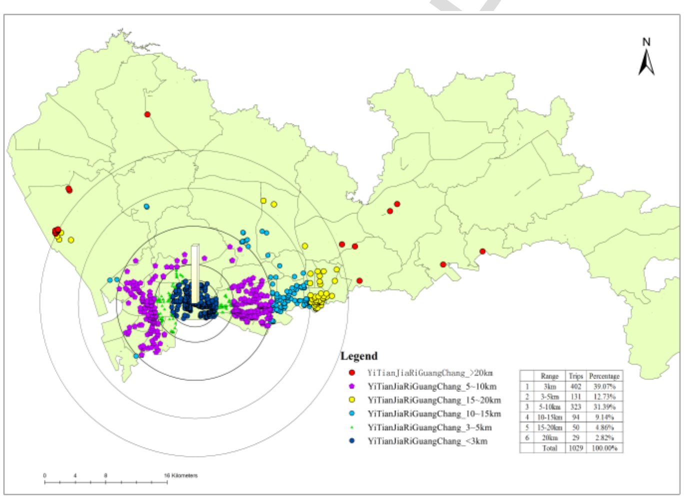
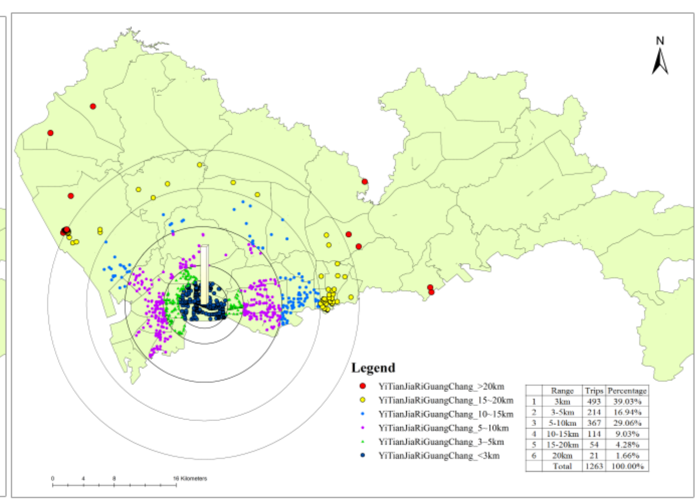
Figure 8. A-degree rankings of 15 CAs at workday

To validate that YiTianJiaRiGuangChang has the highest attractiveness degree among all the commercial centers, revenues and customer satisfaction survey results of this commercial centers are obtained. Total revenue of this commercial center is 1.54 billion in 2014, which is the top one among all the shopping centers in Nanshan District. The customer satisfaction survey from Shenzhen Retail Business Association shows that it is the best one among all the commercial centers [29]. All these support that YiTianJiaRiGuangChang has the highest attractiveness degree.

Each commercial center's real-estate price, i.e., commercial value, is calculated. It is highly related to the area size of the commercial center. YiTianJiaRiGuangChang has an area of 30688.07 m^2 with the ranking of 14^{th} among all 15 commercial centers, and it has the second lowest commercial value among all the 15 commercial center. With such a small size and low commercial value, YiTianJiaRiGuangChang has the highest attractiveness degree among all 15 commercial centers. This shows that attractiveness degree is not necessarily related to the size or the commercial value of the place.

Weekday: Drop-off's origins; Weekday: Pick-up's destinations

Weekend: Drop-off's origins; Weekend: Pick-up's destinations

Figure 9. Taxi's pick-up destinations and drop-off origins of YiTianJiaRiGuangChang for weekday and weekend

YiTianJiaRiGuangChang has the highest attractiveness degree among all 15 commercial centers. To study its attractiveness range, attractiveness thematic maps for YiTianJiaRiGuangChang are drawn for all the pick-ups and drop-offs of taxis in Figure 9. All the taxi visit data of YiTianJiaRuGuangChang are reviewed. All the locations of origins or destinations are covered by 5 circles with the radius of 3km, 5km, 10km, 15km and 20km respectively in Figure 9.

Figure 9 shows that about 90% visits of either drop-offs or pick-ups are from the west (Nanshan district) and the east (Futian district) of YiTianJiaRiGuangChang. This commercial center exactly lies on the boundary of these two districts. It is able to attract customers from both sides. HaiShangShiJie has the lowest attractiveness degree among all 15 commercial centers. Its attractiveness thematic map is shown in Fig. 10.

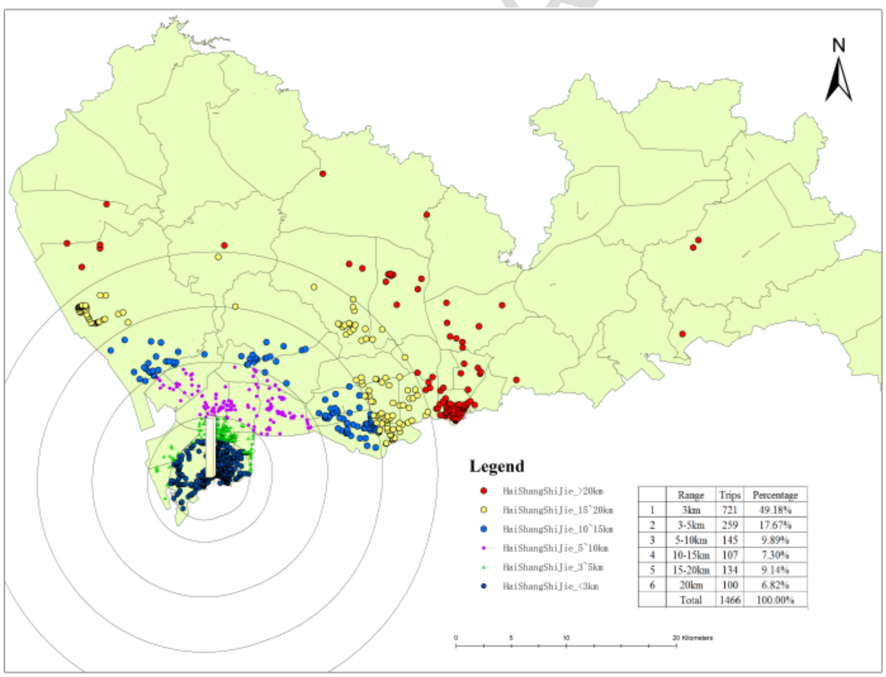


Figure 10. Drop-off origins of HaiShangShiJie on weekend

For YiTianJiaRiGuangChang's taxi pick-ups or drop-offs on weekday or weekend, over 50% visits are within the range of 5 km in Figure 9. For HaiShangSheJie's taxi drop-offs on weekend, about 67% visits are in the 5 km range in Figure 10. The results show that about 50-65% visits by taxi are within a 5 km range. It means most of the customers are the residents nearby and the attractiveness of HaiShangSheJie is much smaller than that of others. In fact, this phenomenon shows that the residential and travel habits for urban customers at Shenzhen visiting commercial centers are much different from that of urban customers from other cities in US or Europe. Urban customers in China are normally densely inhabited and the attractiveness to them mainly depends on the distance to these centers, so that 5 km may be the most popular distance.

VII. Summary and Conclusion

Commercial centers are designed for shopping, recreation, working, and habitation etc. Although customers may live in one or visit one for years, they usually do not know its attractiveness factors behind. Traditional studies are conducted using data from field research, which has a high cost of manpower and time. To overcome this shortage, this study discusses urban commercial centers' attractiveness of taxi travelers based on massive data analytics. All the data are collected from the traffic system, the Internet and the crowd sensing system built in a Cyber-Physical-Social system. Taking Shenzhen as a study case, commercial centers' attractiveness of taxi travelers is evaluated using data statistical analysis and an attractiveness degree model.

Results show that KeJiYuan has the most taxi visits on an average weekday, while Shenzhen University has the most taxi visits on an average weekend day. Commercial centers attract taxi visitors in different periods of time. For example, XinHeChunTian has many visitors, who leave the area after 9pm. This shows that XinHeChunTian has a major recreational function in the evening. HaiShangSheJie has the highest taxi cost (including travel fare, travel time, and travel distance) among all commercial centers. All commercial centers are classified into four categories: long-term-stable, sporadic-burst, day-night-contrast, and emerging-comprehensive type based on the data statistical analysis.

An attractiveness degree model is proposed and used. Results show that YiTianJiaRiGuangChang has the highest attractiveness degree among all the commercial centers. However, this commercial center is with a smaller size and lower commercial value. It is concluded that attractiveness degree is not necessarily related to the size or commercial value of the place, but mainly depends on its locations and services. Attractiveness thematic maps also show that about 50-65% visits by taxi are within a 5 km range.

This is the first time to discuss the attractiveness of commercial centers by integrating massive data from population, area, traffic, and roadmap etc. The results could be applied to market analysis, urban planning, transportation management and related areas.

In the future, our data sources will be extended from taxis to other public transportation. Some sample data have been collected from public transportation for years, such as buses and metros at Shenzhen. However, the problem of diverted traffics for bus routes and metros has not been overcome until now. It is still a challenging problem. One of our plans is to consider them in the future work. Future research can be carried out to consider commercial centers' attractiveness from transit riders or private car users. This topic can also be studied from an evolution view and by identifying fine-grained attractiveness factors.

References

[1] W. Christaller, Die zentralen Orte in Süddeutschland. Jena: Gustav Fischer (Translated by Carlisle W. Baskin, as Central Places in Southern Germany. Englewood Cliffs, NJ: Prentice Hall, 1966), 1933.

[2] C. Dennis, D. Marsland, and T. Cockett. Central place practice: shopping centre attractiveness measures, hinter land boundaries and the UK retail hierarchy. Journal of Retailing and Consumer Services, 9, 2002.

[3] D. L. Huff, A Probabilistic Analysis of Shopping Center Trade Areas Land Economics, vol.39, No. 1, pp. 81–90, Feb. 1963.

[4] W.J. Reilly, The law of retail gravitation. New York: Knickerbocker Press, 1931.

[5] D. McFadden, Economic choices. American Economic Review, vol.91, pp. 351–378, 2001.

[6] V. Severin, Jordan. J. Louviere, A. Finn. The stability of retail shopping choices over time and across countries. Journal of Retailing, vol. 77, pp. 185–202, Summer 2001.

[7] A.G. Wilson. Entropy in urban and regional modeling. London: Pion, 1970.

[8] F. Simini, M. C. Gonza´lez, A. Maritan and A. Baraba´si. A universal model for mobility and migration patterns. Nature, vol.484, pp.96–100, Apr. 2012.

[9] Y. H. Ren, M. Ercsey-Ravasz, P. Wang, M. C. Gonza´lezand Z. Toroczkai. Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. Nature Communications, pp. 1–9, Nov. 2014.

[10] C.E. Dennis, D. W. Marsland, A. Cockett. Why do people shop where they do? Recent Advances in Retailing and Services Science, Proceedings of the 6th International Conference, the European Institute of Retailing and Services Studies, Eindhoven, the Netherlands, 1999.

[11] D. Ashbrook and T. Starner, "Using GPS to Learn Significant Locations and Predict Movement across Multiple Users". Personal and Ubiquitous Computing, Vol.7 (5), pp.275–286, 2003.

[12] Y. Ge, C. Liu, H. Xiong, and J. Chen, A taxi business intelligence system. In Proc. KDD ’11, pages 735–738, 2011.

[13] J. Yuan, Y. Zheng, L. Zhang, X. Xie, and G. Sun. Where to find my next passenger. In Proc. Ubicomp ’11, pages 109–118, 2011.

[14] Y. Zheng, Y. Liu, J. Yuan, and X. Xie. Urban computing with taxicabs. In Proc. Ubicomp ’11, pages 89–98, 2011.

[15] J. Guan and H. Zhou. Research on Space-time Characteristics and Reasons of Idle Taxi Based on FCD: Case in Shenzhen Guomao CBD. ACTA SCIENTIARUM NATURALIUM UNIVERSITATIS SUNYATSENI, vol. 49, pp. 29–36, June, 2010.

[16] J. Yuan, Y. Zheng and X, Xie, "Discovering Regions of Different Functions in a City Using Human Mobility and POIs" Proceedings of ACM Sig KDD Conference on Knowledge Discovery & Data Mining(ACM KDD), pp.186-194, 2012.

[17] G. Qi, X. Li, S. Li, G. Pan, Z. Wang, and D. Zhang. Measuring social functions of city regions from large-scale taxi behaviors. In IEEE PERCOM Workshops, pages 384–388, 2011.

[18] T. Kawasaki and K. W. Axhausen. Choice set generation from gps data set for grocery shopping location choice modelling in canton zurich: comparison to swiss microcensus 2005. Technical report, Swisss Federal Institute of Technology Zurich, 2009.

[19] Y. Yue, H. D. Wang, B. Hu and Q. Q. Li "Identifying Shopping Center Attractiveness Using Taxi Trajectory Data," TDMA'11, September 18, 2011, Beijing, China.

[20] Y. Yue, H. D. Wang, Q. Q. Li, and G. O. Yeh, "Exploratory calibration of a spatial interaction model using taxi GPS trajectories," Computers, Environment and Urban Systems, Vol.36 (2).pp.140–153. Mar. 2012.

[21] Transportation Committee of Wuhan City, Wuhan taxi report, http:// www.whjt.gov.cn/.

[22] Overcrowding commercial centers are facing challenges, http://www.cq.xinhuanet.com/2014-09/15/c_1112486968.htm.

[23] Why is well-designed HuaNan Mall unpopular? http://szbbs.sznews.com/thread-2670804-1-1.html, March, 2015.

[24] OpenStreetMap, http://www.openstreetmap.org/.

[25] WorldPop, http://www.worldpop.org.uk/.

[26] Real Estate Price Map, http://www.szpl.gov.cn/.

[27] Apache Spark, http://spark.apache.org/.

[28] Shenzhen Science & Industry Park(KeJiYuan), http://www.ssipc.com.cn/ web/category.aspx?NodeID=114

[29] Shenzhen Retail Business Association, http://www.szrba.org/.

[30] Wentao Wang, Bo Tang, Xiaopeng Fan, Haixia Mao, Hang Yang, Min Zhu, Efficient Visibility Analysis for Massive Observers, Proceedings of the 8th International Conference on Advances in Information Technology(IAIT2016), December 19~22, Macau China.

[31] Luyan Xiao, Xiaopeng Fan, Haixia Mao, Chengzhong Xu, Ping Lu, Shengmei Luo, When Taxi Meets Bus: Night Bus Stop Planning over Large-scale Traffic Data, Proceedings of the 7th International Conference on Cloud Computing and Big Data (CCBD2016), November 16~18, Macau, China.

[32] Manhua Jiang, Xiaopeng Fan, Fan Zhang, Chengzhong Xu, Haixia Mao, Renkai Liu, Characterizing On-Bus WiFi Passenger Behaviors by Approximate Search and Cluster Analysis, Proceedings of the 7th International Conference on Cloud Computing and Big Data (CCBD2016), November 16~18, Macau, China.

[33] Jinping Guan, Chengzhong Xu, Xiaopeng Fan, Travel Mode Choice Suppression in Large-scale Residential Areas on the City Periphery (Case Study of Jinhexincheng, Shanghai, China), The Transportation Research Board (TRB) 96th Annual Meeting (TRB2017), January 8~12, 2017, Washington, D.C, USA.

[34] J. Guan and D. Yang. Travel Behavior of Two Major Groups in Large-Scale Residential Areas in the Periphery of Shanghai, China: Case Study of Jinhexincheng, Jiading District. Transportation Research Board 92nd Annual Meeting, Washington, D.C., No. 13-0510. 2013.

[35] J. Guan and D. Yang. Residents' Characteristics and Transport Policy Analysis in Large-Scale Residential Areas on a City Periphery: Case Study of Jinhexincheng, Shanghai, China. In Transportation Research Record: Journal of the Transportation Research Board, No. 2512, Transportation Research Board of the National Academies, Washington, D.C., 2015, pp. 11–21.

[36] Huang, Lian, Qingquan Li, and Yang Yue. "Activity identification from GPS trajectories using spatial temporal POIs' attractiveness." Proceedings of the 2nd ACM SIGSPATIAL International Workshop on location based social networks. ACM, 2010.

[37] Giannotti, F., Nanni, M., Pinelli, F., and Pedreschi, D., Trajectory pattern mining. In Proc. of the 13th ACM Intl. Conf. on Knowledge Discovery and Data Mining, SIGKDD, pages 330–339, New York, NY, USA, 2007.

[38] Giannotti, F., Nanni, M., Pedreschi, D., Pinelli, F., Renso, C., Rinzivillo, S., and Trasarti,R. Unveiling the complexity of human mobility by querying and mining massive trajectory data. The VLDB Journal, 20(5):695–719, 2011.

[39] Wei, Ling-Yin, et al. "Pats: A framework of pattern-aware trajectory search." Mobile Data Management (MDM), 2010 Eleventh International Conference on. IEEE, 2010.

[40] Dolega L, Pavlis M, Singleton A. Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. Journal of Retailing and Consumer Services, 2016, 28: 78-90.

[41] Ortegón-Cortázar L, Royo-Vela M. Attraction factors of shopping centers: Effects of design and eco-natural environment on intention to visit. European Journal of Management and Business Economics, 2017, 26(2): 199-219.

[42] Mittal A, Jhamb D. Determinants of shopping mall attractiveness: The Indian context. Procedia Economics and Finance, 2016, 37: 386-390.

[43] A.S. Furtado, R. fileto，C.Renso，M-Attract: Assessing the Attractiveness of Places by using Moving Objects Trajectories Data, GEOINFO 2012 Brazilian Conference on Geographical Information Systems, November 25~27, 2012.

Customer Attractiveness Evaluation and Classification of Urban Commercial Centers by Crowd Intelligence

ACKNOWLEDGMENTS

This research is partially supported by Major Project of Chinese National Programs for Fundamental Research and Development (2015CB352400), National Natural Science Foundation of China (61672513), and Shenzhen Strategic Emerging Industry Development Funds (JCYJ20170818163026031, JSGG20160229200957727).

Customer Attractiveness Evaluation and Classification of Urban Commercial Centers by Crowd Intelligence

Research Highlights:

- Commercial center is precisely defined geographically and sociologically;
- Attractiveness degree is the proposed metric to evaluate the attractiveness;
- A Cyber-Physical-Social System is implemented for massive data processing;
- Statistical analysis and thematic maps shed new insights into attractiveness;
- Commercial centers are classified into four representative types in China.