Research Article

Zhao Yannan, Zhang Lu*, and Zhang Xinhuan*

The spatial distribution of retail outlets in Urumqi: The application of points of interest

https://doi.org/10.1515/geo-2020-0149
received May 17, 2020; accepted October 21, 2020

Abstract: Studies on the distribution of retail outlets are crucial for optimizing urban resource allocation, and their rationalized layout plays an important role in developing urban economies and meeting consumer demand. However, the literature on the subject has been limited by data collection. In the era of big data, there has been an emergence of geo-referenced data that are more accurate and convenient and thus more able to be applied in the retail analysis. This study addresses this lacuna by exploring the spatial distribution of various retail outlets with the application of points of interest. Our study demonstrates that (1) the retail outlets in Urumqi are concentrated in the inner city. The inner-city area has a higher density distribution of retail outlets, and the development of the retail function is more mature. (2) Various categories of retail outlets display dissimilar agglomeration characteristics and hot spots. Specialty stores, clothing and footwear stores, convenience stores, and home-building material markets have lower average nearest-neighbor distances. (3) The retail outlets in the subject area are directionally distributed in the northwest–southeast direction, which is the result of both the topographical condition and government intervention in Urumqi. (4) Various categories of retail outlets tend to agglomerate at different scales. These results are conducive to the exploration of the location rules and layout preferences of retail locations, and they provide a reference for guiding the adjustment and optimization of retail layouts.

Keywords: retail outlets, category, spatial distribution, POIs, Urumqi

*** Corresponding author: Zhang Lu,** Department of International Economics and Trades, School of International Economics and Management, Beijing Technology and Business University, Beijing, 100048, China, e-mail: zhanglu113@mails.ucas.ac.cn
*** Corresponding author: Zhang Xinhuan,** State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China, e-mail: zhangxh@ms.xjb.ac.cn
Zhao Yannan: Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Chaoyang District, Beijing 100101, China, e-mail: zhaoyn.16b@igsnrr.ac.cn

1 Introduction

Rapid urbanization has brought drastic changes to the development of contemporary retail [1]. Although there are lots of online shopping stores, retail stores occupy an important place in modern cities, particularly in terms of economic growth and employment [2,3]. Retail activity, which is one of the most important cost factors for urban residents, has played a powerfully influential role not only in socioeconomic development but also in affecting urban residents' life quality [4–6]. Studies on the spatial distribution of retail outlets have significant importance for land management and the optimization of urban resource allocation [5], and its rationalized layout is critical for the development of urban economies and meeting consumer demands [7,8].

With the continuous development of the retail industry, the issues of retail outlets are gaining increasing attention [9]. Previous studies have focused on retail location [10], hierarchy [11], consumer behavior and commercial space [12], and retail outlet aggregation and competition [13]. Studies on the distribution patterns of retail outlets are of great significance to understanding retail industry rules [14]. However, what has been missed in previous studies is that the retail outlets can be divided into several categories (e.g., supermarkets and convenience stores), which raises the research question that what are the distribution characteristics of various retail outlets. Ignoring the differences is unfavorable to display different characteristics in spatial distribution. One of the reasons for the research lacuna is the lack of data. The data used in retail studies have traditionally mostly been obtained through economic census data and questionnaire surveys, which has led to a lack of timeliness and comprehensiveness of such studies. Generally, two obstacles make it a challenge to obtain proper data. The first obstacle is the limitation of data availability. The fact that the retail industry generates an enormous amount of information makes it very difficult to obtain. The second obstacle pertains to the fact that most information on the retail industry is not available to the public. As noted by Tomlan [15], although commercial activity is everywhere and well-known, the identification and utilization of "good" data are relatively new developments.

Big data bring opportunities and challenges to spatial data mining. In addition to the use of statistical data, and other types of data, open-access data play an increasing role in related studies [16]. Recent advances in some geospatial big data, such as location-based service data [17], points of interest (POIs) [18], and open data from Internet, have been applied to the research of urban structure. Specifically, POIs, i.e., specific point locations that may be useful or interesting, represent an explosion of rich, spatially embedded information. As a consequence, there is a proliferation of research using POIs in commercial studies [19]. In contrast to traditional statistical data, POI data sets represent a much finer grained picture of retail distribution [20]. Gaode map, which is a map software used as a proxy for retail outlets, records the location information and helps to identify spatial distribution characteristics for retail centers. In addition, studies on retail outlets in northwestern China, which has been characterized as a less developed area, are quite limited. Thus, the main aim of our analysis is to analyze the distribution characteristics of various retail outlets with the application of POIs.

The remainder of this study is organized as follows. Section 2 reviews the rationality of the application in retail studies. Section 3 describes the study area, data sources, and methods. Section 4 reports the main results of the study while Section 5 discusses the main findings as well as some limitations of the analysis. Finally, Section 6 presents our main conclusions and some avenues for future research.

2 POIs as data sets in the literature

In the era of big data, we have seen the emergence of georeferenced data which have increasingly attracted interest from geographic information system (GIS) scholars and have become popular in geographical research [21,22]. Increasing numbers of data on POIs are being generated from online platforms [23]. Thus, they are becoming more useful and straightforward for researchers to study properties in space. An enormous need is felt to utilize POIs for retail studies as the pace at which new communities have grown and the sheer number of geospatial variables have been generated in the process cannot be efficiently captured traditionally. In terms of research, we propose POI as an alternative data source for characterizing spatial distribution for the following reasons. The greatest advantage of POIs is that they are entities of interest with well-defined locations. Another obvious feature of POIs is that they are available from service providers as well as in online maps. In addition, POIs have great potential for global analyses since they provide nearly global coverage. With all these advantages, POIs offer new avenues for exploring the geographies of user-generated content.

Existing studies on the application of POIs in retail outlet analysis focused on the three research agendas. The first research agenda is to explore the spatial distribution pattern of retail outlets. The rational distribution of retail outlets is of great significance to the economic development of a city and to satisfy the daily consumption of residents. Some scholars explored the agglomeration and spatial patterns of retail outlets [24]. In addition, previous studies show that the spatial distribution of retail outlets is related to the transport infrastructure [2] and street centrality [20]. The second research agenda focuses on the identification of retail hot spots and the delineation of type zones. The identification of retail outlets plays an important role in the study of urban commercial structure. Previous studies explored the scope, spatial structure, and mode of commercial centers in cities such as Beijing and Guangzhou [25,26]. The third research agenda is to evaluate commercial services which are essential to increase the attractiveness of customers. Previous studies show that there is a spatial distribution difference in customer satisfaction. Among these, the location of metro stations and business centers is the key factor for customers' satisfaction with commercial facilities [27]. Some scholars also discussed the rationality of commercial service facilities' layout [28].

This is the first research agenda to be addressed by our analysis. The overall distribution of retail outlets has been well examined in Western European and Southeast Asian countries and Japan from the perspectives of central place and economic agglomeration theories [29–31]. China has undergone a very dramatic shift from the scarcity of basic consumer goods to relative abundance in recent years. As a country experiencing a transition to a market-oriented economy, Chinese cities provide interesting cases for studying the spatial structure of shopping center. The existing research mainly involves the development mode of shopping centers and mechanisms [32]. It is noteworthy that the market-oriented economy in China has also promoted the explosion of retail outlets, for instance, leisure, entertainment, catering, etc. Thus, the current studies are not conducive for reflecting on the distribution of retail outlets by different categories. The continual emergence of new types of retail business calls for the analysis of how each category of retail outlets distributes. Along with the development of quantitative geography and GISs, mathematical models (e.g., analytical hierarchy process, analysis of correlation, and geographically weighted regression) and spatial statistical analyses (e.g., kernel density estimation, spatial gravity, and hot spot analysis) are new tools for the quantitative analysis of shopping centers [21,33]. Results showed that retail outlets are not evenly distributed in urban spaces and tend to be geographically concentrated. Specifically, homogeneous agglomeration occurs when several retailers selling similar products gather together while heterogeneous agglomeration occurs when many retailers selling different kinds of products [34,35]. In sum, studies on the agglomeration characteristics of different retail enterprises can provide background information for urban planning and development.

3 Methodology

3.1 Study area

From the perspective of location, Urumqi is considered as the "Pivot of Asia" and a "new center of gravity" between Europe, the Middle East, and Central Asia [36]. From the perspective of position, Urumqi is the second largest populous city in northwestern China (Figure 1a). As it is at the intersection of international trade routes (e.g., the Old Silk Road, the Belt, and Road), it has an increasingly active commercial life with varied retail outlets ranging from street shopping to its world-famous covered bazaars [37]. Additionally, Urumqi has great potential for the development of various retail outlets due to its rapid population and income growth. In the past 10 years, the urban population and retail consumption have greatly increased in Urumqi. The urban resident population in Urumqi in 2019 was 3.55 million, which is 1.78 times the number in 2006, and the per capita income was 42,667 yuan in 2019, which is 4.12 times that in 2006 [38]. As a result, the number of various retail outlets has rapidly increased with the addition of various international brands and leisure patterns to meet the needs of emerging retail activities. The total retail sales of consumer goods in 2019 were 138.92 billion, which is 5.05 times that in 2006 [39]. Thus, the special location and position of Urumqi make such research necessary to consider the retail industry in a new light with interesting implications for other cities.

Besides, recent changes in urban planning and the continuous rising retail outlets also indicated the necessity for preliminary exploration of the spatial distribution of retail outlets in Urumqi. Specifically, the spatial distribution of retail outlets in Urumqi is closely related to the urban development and construction process, which is encouraged by the restructuring of the urban space. In the 1980s, the urban construction in Urumqi was dominated by the reconstruction of the old city, and most of the retail outlets were concentrated in the old city. Therefore, the spatial distribution of retail outlets in Urumqi was characterized by a high concentration in the old city. Since the 1990s, the urban construction of Urumqi shifted from the reconstruction of the old city to the development of the new district. The construction of new districts (e.g., the Xinshi District, the Economic Technological District, and the Midong District) witnessed the process of the extension and spreading of new retail outlets to the periphery (Figure 1b). However, retail outlets in the inner city are still in a leading position in terms of quantity, scale, and service level [40].

The continuous improvement in Urumqi's nationalization status and the increasing urban population have brought development opportunities for the retail industry. According to "Urumqi Urban Master Plan 2014–2020," Urumqi is located as an important central city in northwestern China and an international business center to Central and West Asia. Accordingly, relying on surrounding mountains and urban agricultural landscape belts, the land layout pattern of "one axis, two hearts, and seven groups" will be finally formed (Figure 1c). Under this background, studies on the spatial structure characteristics of the retail outlets in Urumqi will help to provide a reference for its spatial layout and structural optimization. Therefore, our analysis attempts to address this research lacuna by exploring the distribution characteristics of retail outlets in Urumqi.

3.2 Data sources

Our study focuses on the central urban area with reference to the "Urumqi Urban Master Plan 2014–2020."

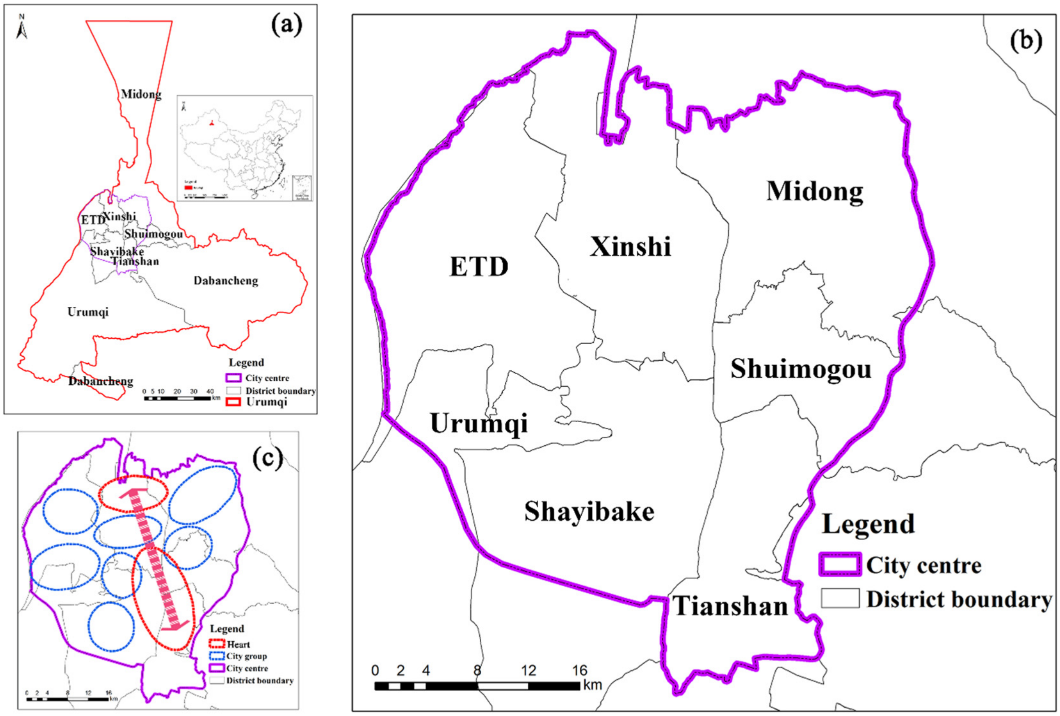

Figure 1: Location of study area.

Three areas, i.e., the downtown area, the three-ring area (Outer and Inner Ring Road, Second Ring Road, and Ring Expressway) and the old city center in the downtown area, are added to the map after the process of registration and vectorization. The properties of retail stores include name, address, longitude, latitude, and business scope, and referred to as POIs applied in our analysis. The amenity POIs data are gathered from Gaode Maps (http://m.amap.com/), which is the primary product of the AutoNavi Software. Given that the positions and properties can be explored through these POIs, after correcting and matching addresses, a total of 17,710 retail centers were extracted. To check the usability of POI data, the information of retail centers are randomly selected and checked by a sampling survey, telephone inquiry, and field visits.

In general, the classification of retail outlets is based on business models, service functions, product structure, location, scale, target customer groups, etc. Regarding the classification standards of retail formats in China, the Chinese national standard "Classification of Retail Formats" (GB/T18106-2000, GB/T18106-2004, GB/T18106-2010) has been issued successively since 2000. Accordingly, the retail centers are divided into 10 categories, which are, namely, convenience stores, grocery stores, supermarkets, discount stores, warehouse member stores, specialty stores, specialty stores, shopping centers, factory direct sales, and department stores. Referring to China's current classification of retail formats (GB/T18106-2010) and taking into account the actual retail industry in Urumqi, this analysis divided the retail centers into nine outlets, including specialty stores, clothing and footwear stores, home-building material markets, convenience stores, shopping malls, appliance electronics stores, supermarkets, integrated markets, and bird and insect markets (Table 1). In total, 4,599 centers in specialty stores (the largest category) and 3,799 centers in clothing and footwear stores were obtained.

The retail outlets in Urumqi are unevenly distributed and do not always match the population density. The population densities of the Tianshan District, the Xinshi District, the Shayibake District, and the Shuimogou District are relatively high, and the number of retail outlets owned by 10,000 people is also relatively high (Table 2). The average density of retail outlets in Urumqi is 1.29/km^2, the average density of the retail outlets in the Tianshan District is 27.58 km^2, and the average density in the Dabancheng District is only 0.01/km^2. Moreover, the number of retail outlets per 1,000 persons in the Tianshan District is 786.52, while the value in Urumqi County is only 84.66, showing a difference of 9.3 times between the two areas.

Table 1: The number and proportion of retail outlets in Urumqi

Outlets	Scope description	Number (parcel)	Proportion (%)
Specialty stores	Specific category of goods, reflecting professionalism, such as office supplies, toy stores, stationery stores, and special trading places (pawn shops, auction houses)	4,599	25.97
Clothing and footwear stores	Mainly clothing, shoes, hats and leather goods stores	3,799	21.45
Home-building material markets	Specializing in the sale of building materials, decoration, and household items in kitchens and bathrooms, such as hardware, steel, and cement stores	2,838	16.02
Convenience stores	Instant products, newspapers and magazines, and daily necessities such as 7 Eleven, Haolinju, and Night and day	3,128	17.66
Shopping malls	Concentrated by a variety of retail stores and service facilities, mainly in the manner of counter sales	1,094	6.18
Appliance electronics stores	Consumer electronics and electrical products such as computers, mobile phones, and digital goods	753	4.25
Supermarkets	Mainly in packaging products such as Carrefour, Aijia.	733	4.14
Integrated markets	Agricultural and sideline products market, fruit market, aquatic seafood market, and vegetable market	584	3.30
Bird and insect markets	Pets, flowers, birds, and fish markets	202	1.14
All retail outlets		17,710	100

3.3 Method

Point-pattern analysis tools discern spatial patterns through the measurement and comparison of the attribute values and geographic location [41]. The spatial point pattern, which consists of a series of points, is driven by an unknown generation mechanism that is irregularly scattered in the study area. Since the distribution of points is quite different, the point mode analysis focuses on the spatial distribution characteristics of all points and their mutual relationship such as aggregation, randomization, and uniform distribution, instead of focusing on single-point attributes. In the study, all types of retail outlets are abstracted as a point, so that a point-pattern analysis can be performed. As a result, the characteristics of human activities and physical processes can be quantified and simulated.

The point-pattern analysis is applied through four steps. First, we use the nearest-neighbor index to describe the spatial agglomeration features of retail centers at different scales. Second, nearest-neighbor hierarchical clustering is applied to analyze the agglomeration characteristics of retail outlets. Third, location entropy analysis is applied to examine the agglomeration features of retail outlets in different categories and circles. Last, we use the method of K function to divide the hierarchy of urban retail business circles in Urumqi.

Table 2: Density of retail outlets in Urumqi

Districts	Population (10,000)	Area (km^2)	Number of outlets	Number of outlet/ 1,000 persons	Number of outlets/km^2	Population/km^2
Tianshan	59.96	171	4,716	786.52	27.58	3,506.43
Shayibake	54.99	422.5	4,150	754.68	9.82	1,301.54
Xinshi	60.72	262.52	4,166	686.10	15.87	2,312.97
Shuimogou	28.74	277.56	2,140	744.61	7.71	1,035.45
ETD Zone	23.73	275.59	1,250	526.76	4.54	861.06
Dabancheng	4.16	4,759.18	52	125.00	0.01	8.74
Midong	28.35	3,407.42	1,183	417.28	0.35	83.20
Urumqi county	6.26	4,141	53	84.66	0.01	15.12
Total	266.91	13,716.77	17,710	663.52	1.29	194.59

Note: ETD Zone refers to the economic technological district.

3.3.1 Nearest-neighbor index

The nearest-neighbor index examines the distances between each point and its closest points [42]. The function of the nearest-neighbor index is calculated by the following equation:

$$\mathrm{NNI} = \frac{r_\alpha}{r_r} = \frac{2\sum_i d_i/N}{\sqrt{N/A}} \tag{1}$$

where r_α refers to the average nearest-neighbor distance between i and its nearest neighbor, r_r refers to the expected random distance, N refers to the number of points in the research area, and α refers to the area of study. The value of the nearest-neighbor index is usually above 0. For an NNI value above 1, the spatial distribution pattern indicates a scattered distribution. For an NNI value below 1, the spatial distribution pattern indicates a clustered distribution. If the NNI = 1, the points in the study area are randomly distributed.

3.3.2 Nearest-neighbor hierarchical clustering

The nearest-neighbor hierarchical clustering algorithm is an algorithm used to detect hot spots in the spatial distribution of feature points. The algorithm aims to find different pairs of clusters and to merge them by following the paths in the nearest-neighbor graph of the clusters [43]. By defining the threshold of distance and comparing it with the distance of each set of point pairs, the point will be counted in the aggregate units when the distance between one point and other points (at least one) is lower than the threshold. As a result, multilevel hot spots will be obtained [44].

3.3.3 Location entropy analysis

Location entropy is a widely used indicator in judging the distribution of a certain spatial element. The application of the analysis is used to reflect the degree of specialization of a certain industry or enterprise [45,46]. This study uses location entropy to examine the specialization of retail outlets located in different categories and circles. The parameter is defined to capture these features and is calculated as follows:

$$N_{K-A} = n_{K-A}/n_K \tag{2}$$

where N_{K-A} refers to the location entropy value of retail outlets A in the Kth tread zone, n_{K-A} refers to the number of retail outlets A in the Kth tread zone, and n_K refers to the total number of retail outlets in the Kth tread zone. Therefore, N_{K-A} indicates the ratio of the number of retail outlets A in the Kth tread zone to the total number of retail outlets in the Kth tread zone. Empirically, the higher the N_{K-A} value, the higher the specialization of retail outlets in the Kth tread zone.

3.3.4 Ripley's *K*

Ripley's K-function [47] and the weighted K-function are popular methods to determine whether a set of point features is clustered at multiple different distances. Since the distribution pattern of point elements varies with the scale, Ripley's K-function is used to reflect the specific distribution mode of point elements at different scales. It is one of the most commonly used methods for point-pattern analysis.

The Ripley's $K(d)$ represents the ratio of the average of the actual point elements to the density of the sample points within the distance d and the formula is as follows:

$$K(d) = A\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{w_{ij}(d)}{n^2}; \tag{3}$$

where n refers to the number of point elements, $w_{ij}(d)$ refers to the distance between i and j of within distance d, and A refers to the area of the research area. Furthermore, Ripley's K function measures whether the spatial distribution pattern of the actual observation point elements is a clustered, scattered, or random distribution. The related indicators are calculated as follows:

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d \tag{4}$$

4 Results

4.1 Distribution characteristics in each tread zone

The circle analysis adopts centers on the geometric center of a specific analysis object. Then concentric circles with different radii and the same spacing are formed. This method has obtained a certain level of consensus and is applied in our analysis. One of the issues in the application of circle analysis is the determination of the radius. In our study, 5 km is chosen as the threshold radius referring to (1) in general, human needs can be resolved mostly within a radius of 5 km and (2) the idea of “5-km life circle” has been widely accepted in urban planning in China. The circle is centered on the People’s Square; and then based on the circle, six circles are expanded outward at the same interval. As a result, all points in the retail outlets are divided into six tread zones in each buffer radius of 5, 10, 15, 20, and 25 km. The six tread zones are named as follows: the primary tread zone (0–5 km), secondary tread zone (5–10 km), third tread zone (10–15 km), fourth tread zone (15–20 km), fifth tread zone (20–25 km), and sixth tread zone (25–30 km), from the inner city to the periphery of the city. The distribution characteristics in each tread zone are shown in Figure 2. The following two initial observations can be made about the spatial distribution characteristics.

(1) In summary, the spatial distribution of retail outlets in Urumqi shows a center directional sign (Figure 2). In terms of the overall number of the retail outlets in each tread zone, the primary tread zone is where most (10,137) of the retail outlets are located, accounting for 57.24% of the total. In other words, the central tread zone is the main distribution area for retail outlets. As the distance increases, the number of retail outlets first decreases and then increases. For instance, there are 5,258 outlets in the second tread zone, 695 in the third tread zone, and 1,248 in the fourth tread zone. Moreover, the number of retail outlets in the fifth tread zone reaches the lowest level, 179 outlets, and the number again increases in the sixth tread zone, with 204 outlets.

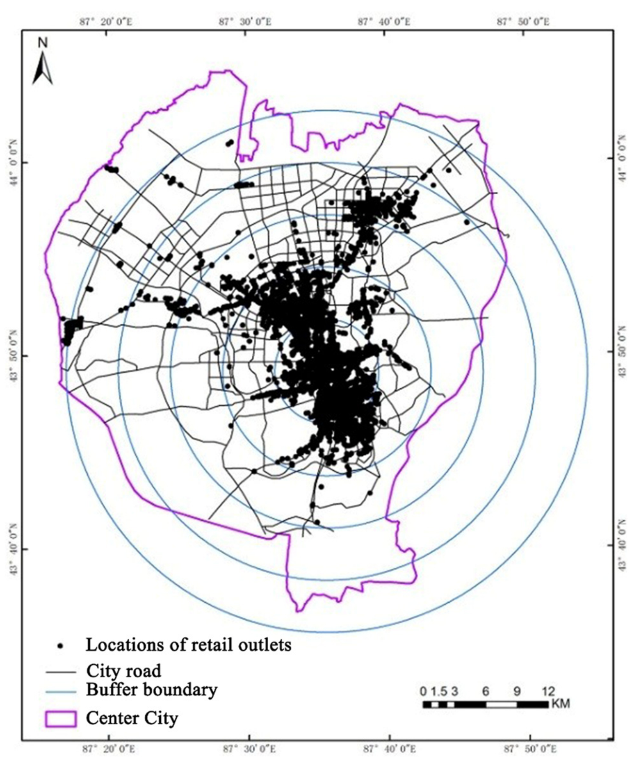

Figure 2: The distribution of overall retail outlets in tread zones.

Comparing the various retail outlets with the outlets in the first tread zone, it is found that the number of convenience stores (Figure 3a), clothing and footwear stores (Figure 3c), specialty stores (Figure 4a), and bird and insect markets (Figure 4b) decreases most in the second tread zone. In the third tread zone, compared with the outlets in the second tread zone, the number of specialty stores (Figure 4a), appliance electronics stores (Figure 4c), home-building material markets (Figure 5a), and integrated markets (Figure 5c) decreases the most. In the fourth tread zone, compared with the outlets in the third tread zone, the number of all outlets increases except shopping malls (Figure 5b), and the number of specialty stores increases the most (Figure 4a). From the fourth tread zone to the sixth tread zone, the outlets including supermarkets (Figure 3b) and home-building material markets (Figure 5a) maintain relatively stable levels. The rate of change in different retail outlets in different spheres shows their preference for location selection. The low rate of increase and the low rate of decrease indicate that their location adaptability is relatively strong and that their sensitivity to location preference is weak.

(2) Each tread zone is dominated by different retail outlets (Figure 6). The observations in terms of each tread zone are as follows. In the primary tread zone, the retail outlets with location entropies larger than 1 are clothing and footwear stores, specialty stores, and bird and insect markets. In the second tread zone, the retail outlets with location entropies larger than 1 are convenience stores, supermarkets, home-building material markets, and shopping malls. The third tread zone is dominated by convenience stores, supermarkets, shopping malls, and integrated markets. Moreover, the retail outlets with location entropies larger than 1 in the fourth tread zone are convenience stores, supermarkets, appliance electronics stores, home-building material markets, and integrated markets. Reviewing the fifth tread zone, it is noteworthy that appliance electronics stores, convenience stores, and supermarkets have high location entropies compared with other retail outlets. In the sixth tread zone, retail outlets including integrated markets, appliance electronics stores, and supermarkets have relatively high location entropy values.

The number of retail outlets in the fifth and sixth zones is 383, accounting for 2% of the total. Thus, only the first four zones are discussed subsequently. In reviewing the first four zones, the following observations

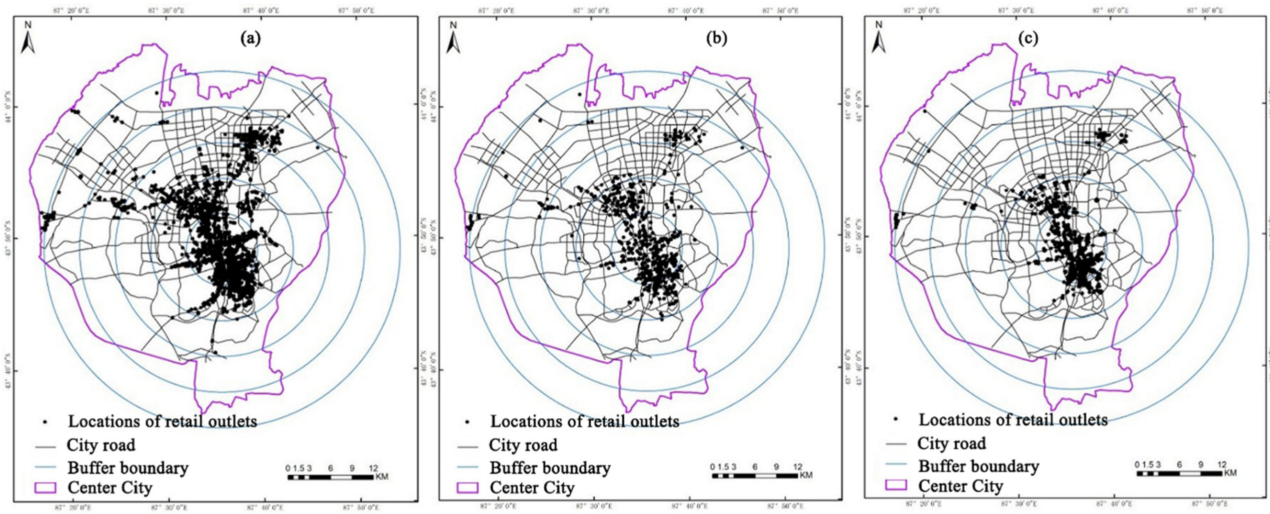

Figure 3: The distribution of retail outlets in tread zones ((a) convenience stores, (b) supermarkets, and (c) clothing and footwear stores).

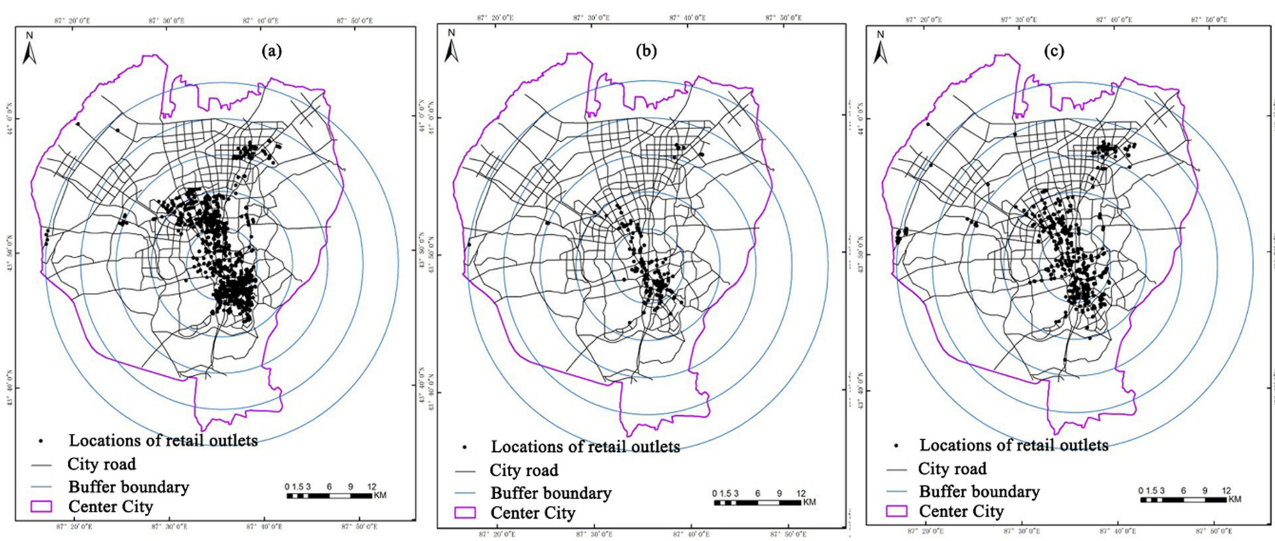

Figure 4: The distribution of retail outlets in tread zones ((a) specialty stores, (b) bird and insect markets, and (c) appliance electronic stores).

are made. The location entropy of convenience stores and supermarkets peaks in the third tread zone (10–15 km) and is the lowest in the primary tread zone (0–5 km). The location entropies of clothing and footwear stores, specialty stores, and bird and insect markets peak in the primary tread zone (0–5 km) and are the lowest in the third tread zone (10–15 km). The location entropy of appliance electronics stores peaks in the fourth tread zone (15–20 km) and is the lowest in the third tread zone (10–15 km). The location entropy of home-building material markets peaks in the second tread zone (5–10 km) and is the lowest in the primary tread zone (0–5 km). The location entropies of shopping malls and integrated markets peak in the third tread zone (10–15 km) and are the lowest in the primary tread zone (0–5 km) (Figure 6).

4.2 The agglomeration characteristics of retail outlets

In summary, retail outlets show significant clustering characteristics. The nearest-neighbor indexes of all retail outlets are less than 1, and the Z-test values are less than −2.58, indicating an obvious aggregative distribution under the 1% significance level (Table 3).

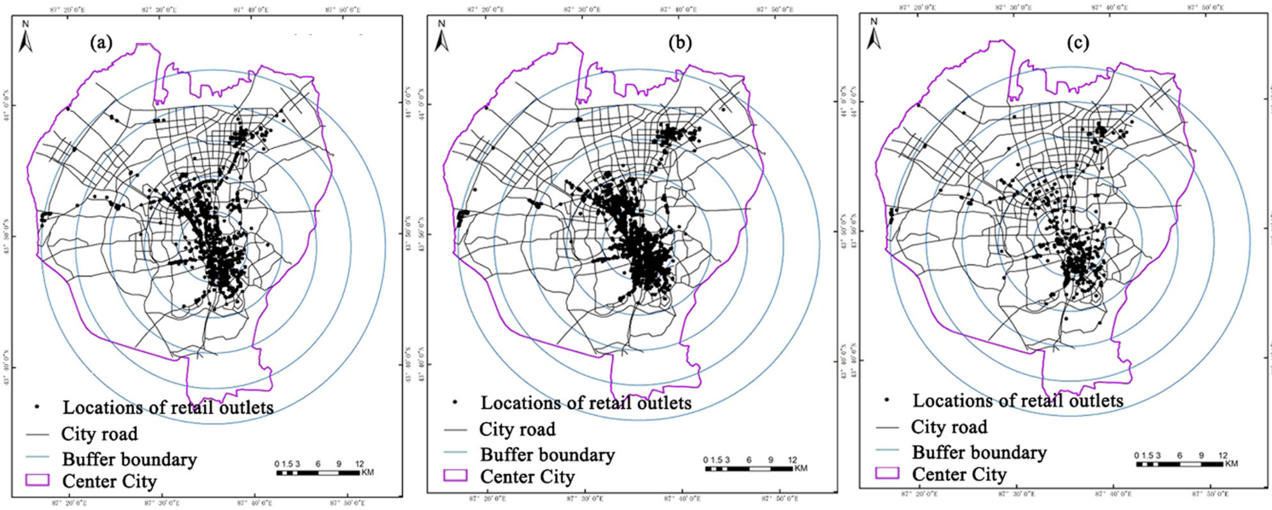

Figure 5: The distribution of retail outlets in tread zones ((a) home-building material markets, (b) shopping malls, and (c) integrated markets).

Furthermore, since significant results were obtained with all the retail outlets (Figure 7), clustering characteristics in different directions can be made with the application of the nearest-neighbor hierarchical clustering algorithm. The significant second-order hot spots can be observed in the retail outlets of supermarkets (Figure 8b), shopping malls (Figure 9a), bird and insect markets (Figure 9b), and appliance electronics stores (Figure 9c; Table 4). Moreover, significant third-order hot spots can be seen in the retail outlets of convenience stores (Figure 8a), clothing and footwear stores (Figure 8c), home-building material markets (Figure 10a), specialty stores (Figure 10b), and integrated markets (Figure 10c).

Specifically, regarding convenience stores, there are 88 first-order hot spots and 6 second-order hot spots (Figure 8a), which are located in satellite plaza, the People's Government, the intersection of West Jinshajiang Road and Karamay Road, the Midong District, east of East Ring Road, and the Black Mountain area. Additionally, there is a third-order hot spot zone between the South of Suzhou Road and north of Outer Ring Road. Regarding the supermarkets, there are 15 first-order hot spots, 10 of which are located in the inner city. Additionally, two second-order hot spots are found in the north of Xinyi Road and the inner city. Regarding the clothing and footwear stores, there are 80 first-order hot spots, eight

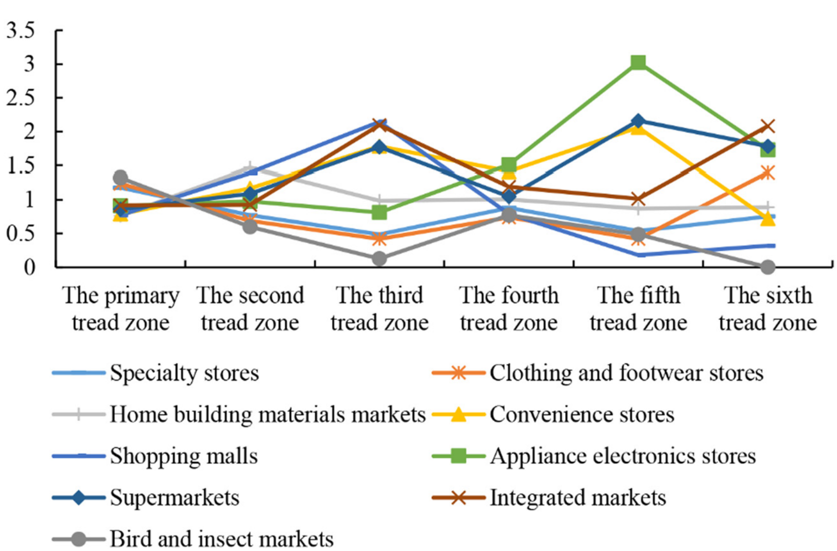

Figure 6: The location entropy of retail outlets in tread zones.

Table 3: The nearest-neighbor hierarchical clustering analysis of retail outlets

Category	Number of outlets	Average nearest-neighbor distance	Expect average nearest-neighbor distance	Nearest-neighbor index	Z value	Significance level (%)
Specialty stores	4,599	45.44	226.84	0.20	−103.75	1
Clothing and footwear stores	3,799	29.60	225.48	0.13	−102.16	1
Convenience stores	3,128	86.96	306.09	0.28	−76.60	1
Home-building material markets	2,838	58.45	291.53	0.20	−81.48	1
Shopping malls	1,094	135.95	418.14	0.33	−42.70	1
Appliance electronics stores	753	138.38	555.83	0.25	−39.43	1
Supermarkets	733	222.59	536.41	0.42	−30.30	1
Integrated markets	584	264.68	624.45	0.42	−26.64	1
Bird and insect markets	202	288.04	929.16	0.31	−18.76	1
All retail outlets	17,710	21.57	131.99	0.16	−212.99	1

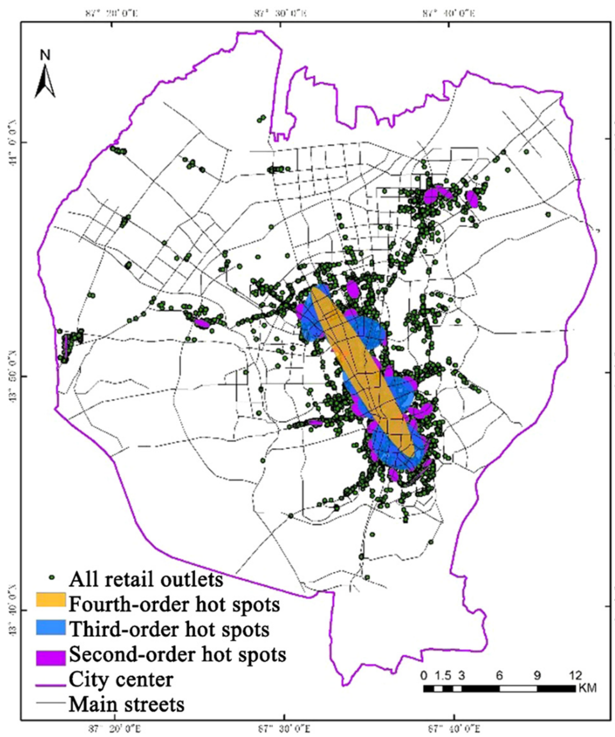

Figure 7: Nearest neighbor index of the overall retail outlets in Urumqi.

second-order hot spots, and 1 third-order hot spot area. Regarding the shopping malls, there are 24 first-order hot spots between the north of Suzhou Road and south of Kashi Road and 2 second-order hot spots in the inner city. Regarding the bird and insect markets, there are 7 first-order and 2 second-order hot spots. Regarding the home-building material markets, 56 first-order, 5 second-order, and 1 third-order hot spots are formed. Regarding the appliance electronics stores and integrated markets, there are 18 first-order hot spots and 1 second-order hot spot in the inner city. Regarding the specialty stores, there are 129 first-order, 11 second-order, and 1 third-order hot spots.

4.3 The agglomeration at different scales

With the application of the *K* function, the spatial characteristics in various retail outlets at different scales can be observed. As a result, the distance range of the location layout for all retail outlets is obtained since the retail outlets are agglomerated at different scales. According to the analysis results, within 12 km, the actual function values in all retail outlets are larger than the upper limit of the confidence interval, and the significance test

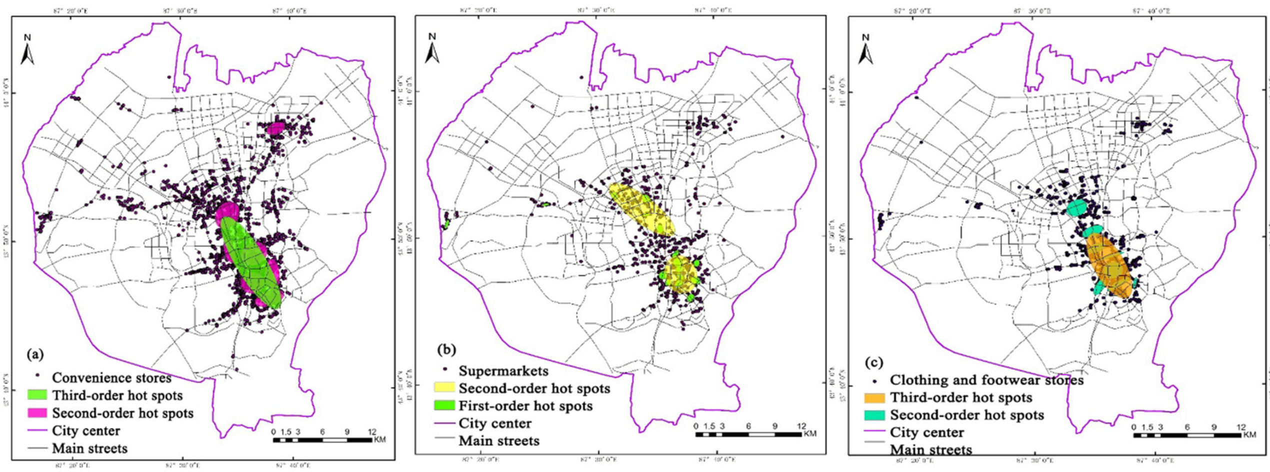

Figure 8: Nearest neighbor index of retail outlets in Urumqi (convenience stores, supermarkets, and clothing and footwear stores).

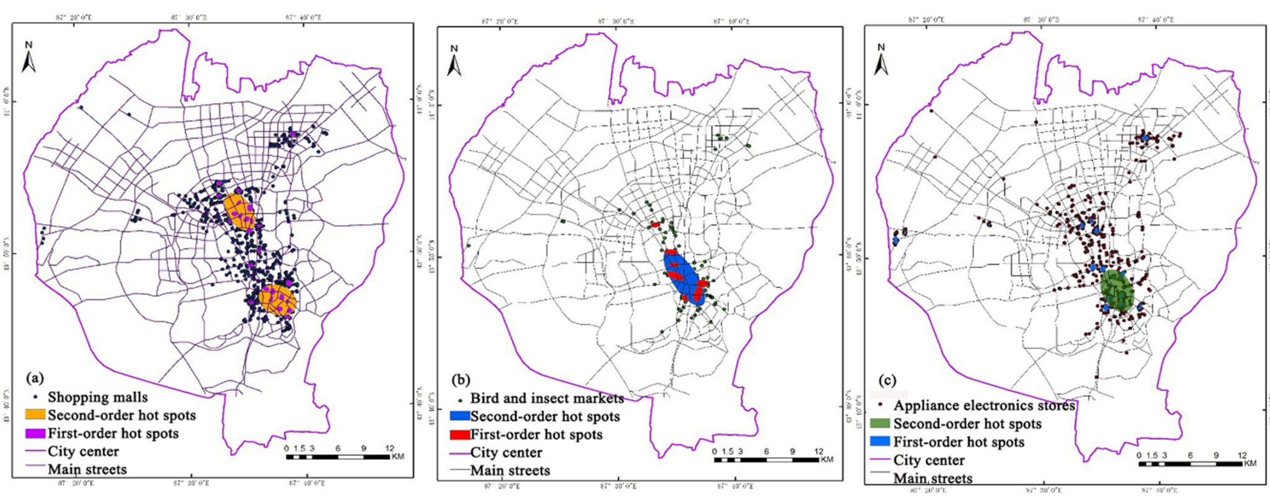

Figure 9: Nearest neighbor index of retail outlets in Urumqi (shopping malls, bird and insect markets, and appliance electronic stores).

passes. In other words, within the observation range of 12 km, all retail outlets show obvious agglomeration characteristics.

While observing the trend of the $L(d)$ function curve, the clustering characteristics of all retail outlets are similar, which presents an inverted U. However, the spatial scales of the retail outlets vary. Regarding the tendency of the L-value curve in each retail outlet, a decreasing trend can be observed.

The appearances of the distance peaks differed. Among them, (1) convenience stores have the largest agglomeration value at a distance of 7.26 km (Figure 11a), which show that the convenience stores exhibit agglomeration characteristics in a larger scale and that the convenience stores can be sited in a larger space. (2) Home-building material markets (Figure 11b) and supermarkets (Figure 11a) also have large spaces for siting, which are agglomerated at distances of 6.22 and 6.20 km, respectively. (3) The peak distances of the retail outlets, such as integrated markets, appliance electronics stores, shopping malls, and specialty stores, decrease, which are 5.51, 5.24, 5.21, and 5.11 km, respectively (Figure 11b). Therefore, these retail outlets have smaller ranges of distance for siting. (4) The clothing and footwear stores and bird and insect markets have the lowest peak distances.

5 Discussion

Retail activity and location are closely related to the economic and social progress of cities. The research

Table 4: The number of the first-order and second-order hot spot in retail outlets

Category	First-order hot spot	Second-order hot spot
Specialty stores	129	11
Convenience stores	88	6
Clothing and footwear stores	80	8
Home-building material markets	56	5
Shopping malls	24	2
Appliance electronics stores	18	1
Integrated markets	18	1
Supermarkets	15	2
Bird and insect markets	7	2
All retail outlets	435	38

presented here suggests methodological approach which relates to a concern for the representativeness of POI data. The distribution of retail outlets in Urumqi displays that the center is favored over the periphery. Retail outlets are more likely distributed in the inner city. This characteristic is consistent with the distribution of the population, in that there is high population density in the inner city. Customers in the inner city tend to have higher buying power than that of the periphery. Our study is consistent with the traffic and market principles that retailers tend to agglomerate in places with good traffic conditions and high population density [48–50]. Moreover, the retail outlets in Urumqi show significant clustering characteristics. From an economic point of view, retail externality is considered one of the important causes of retail agglomeration. Retail agglomeration reduces the uncertainty and search costs for consumers [51–53]. As a result, the agglomerated retail outlets are attractive to consumers and thus increase the overall benefit of retail.

Our analysis shows that each category of retail outlets is agglomerated in space. According to Hotelling's Principle of Minimum Differentiation, retailers opening stores near competitors can expand sales areas and maximize profits [54]. Various categories of retail outlets exhibit different service radii and agglomeration characteristics. Specifically, three retail outlets, including specialty stores, convenience stores, and clothing and footwear stores, have large numbers of first-order hot spots. Four retail outlets, including home-building material markets, shopping malls, appliance electronics stores and integrated markets, have fewer hot spots. Supermarkets are also significantly affected by population density. However, high density of the subject area has made it spatially discretely distributed to achieve large service coverage. Moreover, bird and insect markets have large ranges of services and, therefore, do not exhibit a significant agglomeration. The agglomeration differences in retail outlets are in line with Reilly's law of retail gravitation [55] and Huff's spatial interaction model [56]. Both approaches calculate the power of retailers as being directly proportional to its size and inversely proportional to the distance from the consumer.

In addition, the distribution of retail outlets is also constrained by the topographical condition and urban development strategy in Urumqi. The retail outlets in Urumqi show a directional distribution in the northwest–southeast direction, which is the result of both

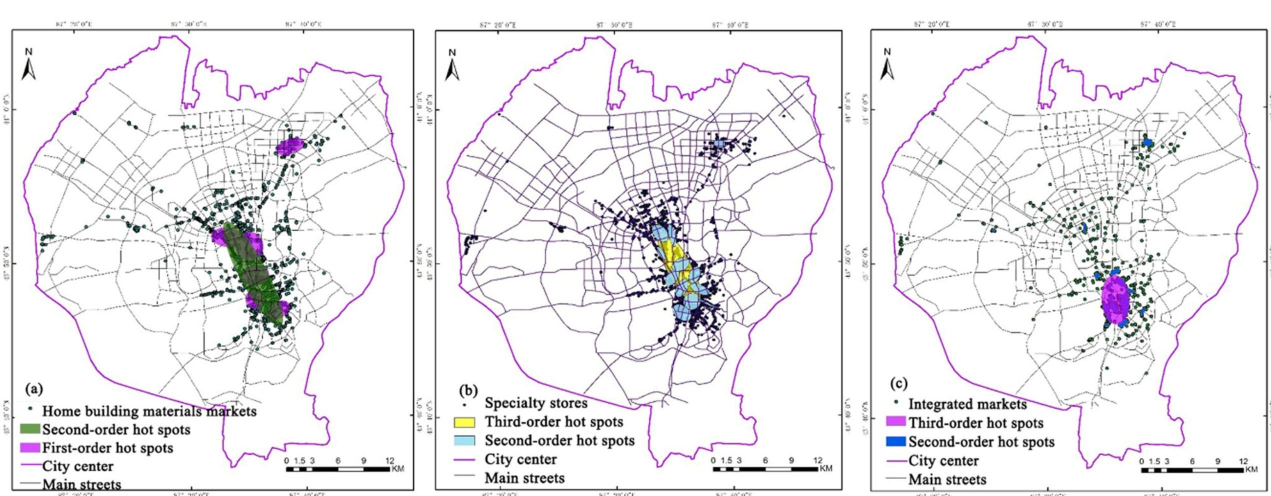

Figure 10: Nearest neighbor index of retail outlets in Urumqi (home-building material markets, specialty stores, and integrated markets).

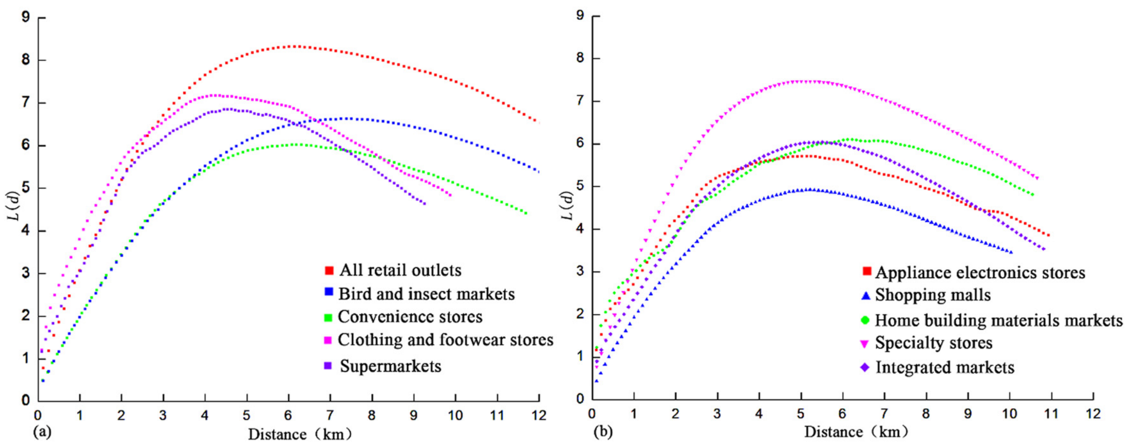

Figure 11: Results of Ripley' *K* on retail outlets in Urumqi.

the topographical condition and government intervention. Specifically, human activity is constrained by the topographical condition, which is bordered on three sides (except the north side) by mountains. Additionally, under the direction of the urban development strategy, which is "controlling southwards and expanding northwards," the population is increasingly agglomerated in the surrounding area of that direction.

Our analysis shows additional potential issues of layout characteristics in Urumqi. For instance, retail outlets compete extensively in the inner city and are unevenly distributed in new districts, which appears to be an unreasonable layout state according to central place theory. As Seim [57] showed, retail merchants must consider not only the location but also the distance from competitors. To prevent these shortcomings, intervention from the outside (e.g., local government) is necessary. Empirical experiences have shown that retail planning policies are very diverse and respond to society's needs and national specificities [58–60].

The current analysis suffers from several limitations, which indicate a potential direction for future research. The first limitation relates to the accuracy of the data obtained with respect to the retail outlets, which must be further improved. The POI data used in this study are obtained from Gaode and Tencent open data. After correction and address matching, although the data are confirmed by a sampling survey, telephone inquiry, and field visit, the data are massive and cannot be fully investigated and confirmed. Thus, data classification deviations might exist. Besides, there is a lack of analysis of factors behind the generation of distributional characteristics. Thus, studies exploring factors such as road network density and building density can be conducted in the future.

6 Conclusion

While the retail activity keeps developing, what remains to be examined is the spatial structure under various retail outlets. The rational layout of retail outlets is of great significance in meeting the consumption needs of urban residents, promoting urban economic development, and optimizing the allocation of land resources. Confined to data availability, the previous studies had difficulties reflecting the spatial layout and development of retail outlets. In recent years, due to the rapid development of big data, data sources for urban space studies are increasingly abundant. This study uses POIs as the data source to identify the distribution of retail outlets in Urumqi. The retail outlets in Urumqi present the following distribution characteristics: (1) the distribution of retail outlets density generally decreases and entropy value increases from the inner city to the periphery, indicating that retail outlets in the inner city are distributed densely and comprehensively. (2) Various categories of retail outlets display dissimilar agglomeration characteristics and hot spots. Specialty stores, clothing and footwear stores, convenience stores, and home-building material markets have lower average nearest-neighbor distances. Significant third-order hot spots can be seen in the retail outlets of convenience stores, clothing and footwear stores, home-building material markets, specialty stores, and integrated markets. (3) Influenced by the topographical condition and urban development strategy, the retail outlets are agglomerated in different directions. (4) Various categories of retail outlets tend to agglomerate at different scales.

We believe our analysis makes contributions to the literature from the following three aspects. (1) The central contribution of our analysis is to set out an alternative data source for spatial analysis of retail outlets. The development of information technologies has provided us with high-volume high-velocity data and also a variety of spatially and temporally referenced data. Grasping the opportunities of the big data era is critical for urban development [33]. In sum, the arrival of the big data era has injected new blood into the study of commercial geography. (2) The second contribution is to remind researchers to differentiate various categories of retail outlets in commercial studies. Relatively limited attention has been paid to explore the various retail outlets categories. One of the advantages of POI data is that it helps to display precise distributions of various retail outlets. In this regard, our analysis offers empirical evidence for understanding the different characteristics among retail outlet categories. By identifying the distribution rules of different retail outlets and summarizing the spatial distribution of retail outlets in different categories, our analysis plays an important role in the development of retail activities and the layout of retail facilities. (3) The third contribution is to enrich the empirical studies on commercial geography. Previous studies on retail outlets mainly focus on the eastern cities in China, for instance, Beijing, Shanghai, Shenzhen, Changchun, etc. Located in the plain area, these cities have relatively better development basis and conditions. In comparison, Urumqi is located in the northwest of China and surrounded by mountains on three sides. The spatial layout of retail outlets in Urumqi has been largely influenced by the terrain. Thus, its spatial distribution of retail outlets in Urumqi is significantly different from that of eastern cities.

The practical significance of this study is that through a detailed analysis of the spatial distribution characteristics of different retail outlets in Urumqi, our study helps provide references for the rational layout and optimization of retail outlets there. Therefore, the practical implications of this study are twofold. First, by investigating the spatial distribution of retail outlets, our study aims to identify a set of candidate areas in the city for merchants to open a store. The geographical placement of a new retail outlet has been of prime importance in trading and commercial systems in cities [61]. Second, our study can provide evidence-based background information for retail planning by reminding policy makers that intervention from local government is necessary to prevent excessive competition; and, thus, it is necessary to employ such intervention accordingly. Specifically, excessive agglomeration of similar retail outlets can cause a negative effect on economic development. This observation shows that although retail agglomeration is a self-organizing choice, this self-organizing process is not entirely sensible. Especially in the context of the rapid urbanization that is prevalent in Chinese cities, self-organization will quickly fail. Therefore, intervention from the outside is necessary. Only a combination of self-organization and retail planning can promote the successful development of the retail industry and city development.

Acknowledgments: This work was financially supported by "Light of West China" Program of CAS(2019-XBQNXZ-A-005) and China Postdoctoral Science Foundation (2020M680660).

Author contributions: Zhao Yannan developed the original idea for the study and conceived of and designed the methodology together. Zhao Yannan drafted the manuscript, which was revised by Zhang Lu, Zhang Xinhuan. All authors have read and approved the final manuscript. We would like to thank the anonymous reviewers and editor for their constructive comments on an earlier version of this paper.

Conflict of interest: The authors declare no conflict of interest.

Data availability statement: Some or all data, models, or code generated or used during the study are available from the first author by request (zhaoyn.16b@igsnrr.ac.cn).

References

[1] Ding CR, Zhao XS. Assessment of urban spatial-growth patterns in China during rapid urbanization. Chin Economy. 2014;44(1):46–71.
[2] Rui Y, Yang Z, Qian T, Khalid S, Xia N, Wang J. Network-constrained and category-based point pattern analysis for Suguo retail stores in Nanjing, China. Int J Geographical Inf Sci. 2016;30(2):186–99.
[3] Kang CD. Spatial access to pedestrians and retail sales in Seoul, Korea. Habitat Int. 2016;57:110–20.
[4] Brown S. Institutional change in retailing: A review and synthesis. Eur J Mark. 1987;21(6):5–36.
[5] Lina H, Bin Z, Mingwu G, editors. Study on the spatial pattern of urban commercial service based on Voronoi model. Management and Service Science, 2009 MASS'09 International Conference on; IEEE: 2009.
[6] Kang C-D. Valuing spatial access to types of retail and effects on the housing price in Seoul, Korea. J Urban Plan Dev. 2018;144(2):05018007.
[7] Goldman A. The transfer of retail formats into developing economies: The example of China. J Retail. 2001;77(2):221–42.
[8] Wang S, Jones K. Retail structure of Beijing. Environ Plan A. 2002;34(10):1785–808.

[9] Reinartz W, Dellaert B, Krafft M, Kumar V, Varadarajan R. Retailing innovations in a globalizing retail market environment. J Retail. 2011;87:S53–S66.
[10] Wang FH, Chen C, Xiu CL, Zhang PY. Location analysis of retail stores in Changchun, China: A street centrality perspective. Cities. 2014;41:54–63.
[11] O hUallachain B, Leslie TF. Spatial pattern and order in sunbelt retailing: Shopping in Phoenix in the twenty-first century. Professional Geographer. 2013;65(3):396–420.
[12] Kumar A, Trivedi M, Bezawada R, Sridhar K. A comparative analysis of differential consumer response across supermarket and specialty store in the candy category. J Retail Consum Serv. 2012;19(6):561–9.
[13] Li Y, Liu L. Assessing the impact of retail location on store performance: A comparison of Wal-Mart and Kmart stores in Cincinnati. Appl Geogr. 2012;32(2):591–600.
[14] Enru W. Understanding the 'retail revolution'in urban China: A survey of retail formats in Beijing. Serv Industries J. 2011;31(2):169–94.
[15] Tomlan MA. Historic Preservation: Caring for Our Expanding Legacy: Springer; 2014.
[16] Lou G, Chen Q, He K, Zhou Y, Shi Z. Using nighttime light data and poi big data to detect the urban centers of hangzhou. Remote Sens. 2019;11(15):1821–40.
[17] Cai JX, Huang B, Song YM. Using multi-source geospatial big data to identify the structure of polycentric cities. Remote Sens Environ. 2017;202:210–21.
[18] Yao Y, Li X, Liu XP, Liu PH, Liang ZT, Zhang JB, Mai K. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. Int J Geographical Inf Sci. 2016;31(4):825–48.
[19] Yang S, Wang M, Wang W, Sun Y, Gao J, Zhang W, Zhang J. Predicting commercial activeness over urban big data. Proc ACM Interactive, Mobile, Wearable Ubiquitous Technol. 2017;1(3):119.
[20] Lin G, Chen XX, Liang YT. The location of retail stores and street centrality in Guangzhou, China. Appl Geogr. 2018;100:12–20.
[21] Graham M, Shelton T. Geography and the future of big data, big data and the future of geography. Dial Human Geograp. 2013;3(3):255–61.
[22] Miller HJ, Goodchild MF. Data-driven geography. Geo J. 2015;80(4):449–61.
[23] Jiang S, Alves A, Rodrigues F, Ferreira J, Pereira FC. Mining point-of-interest data from social networks for urban land use classification and disaggregation. Computers, Environ Urban Syst. 2015;53:36–46.
[24] Sevtsuk A. Location and agglomeration: The distribution of retail and food businesses in dense urban environments. J Plan Educ Res. 2014;34(4):374–93.
[25] Fang W, Xiaolu G, Zening X. Identification and classification of urban commercial districts at block scale. Geographical Res. 2015;34(06):1125–34.
[26] Chen W, Liu L, Liang Y. Retail center recognition and spatial aggregating feature analysis of retail formats in Guangzhou based on POI data. Geographical Res. 2016;35:703–16.
[27] Teng W, Yandong W, Xiaoming Z, Xiaokang F. Spatial distribution pattern of the customer count and satisfaction of commercial facilities based on social network review data in Beijing, China. Computers, Environ Urban Syst. 2018;71:88–97.
[28] Zhou XX, Ding Y, Wu CB, Huang J, Hu CD. Measuring the spatial allocation rationality of service facilities of residential areas based on internet map and location-based service data. Sustainability. 2019;11(5):1337.
[29] Dawson J. Market centers and retail location. Urban Stud. 1991;28(3):497–9.
[30] Borgers A, Vosters C. Assessing preferences for mega shopping centres: A conjoint measurement approach. J Retail Consum Serv. 2011;18(4):322–32.
[31] Sahara T, Watanabe T. A study of retail format strategy deployed by Southeast Asia regional retailers: Focusing on cases in Vietnam. J Mark Distrib. 2016;18(2):77–99.
[32] Shi Y, Wu J, Wang S. Spatio-temporal features and the dynamic mechanism of shopping center expansion in Shanghai. Appl Geogr. 2015;65:93–108.
[33] Kitchin R. Big data and human geography: Opportunities, challenges and risks. Dialogues Hum Geogr. 2013;3(3):262–7.
[34] Hanson JW. A proposed paradigm for consumer product disposition processes. J Consum Aff. 1980;14(1):49–67.
[35] O'Kelly ME. A model of the demand for retail facilities, incorporating multistop, multipurpose trips. Geographical Anal. 1981;13(2):134–48.
[36] Dong W, Zhang XL. Urumqi. Cities. 2011;28(1):115–25.
[37] Zhao Y, Yang D, Zhang X, Xiong C, Chen D. Temporal and spatial evolution of top grade hotel industry and its influencing factors in central district of Urumqi. J Univ Chin Acad Sci. 2017;34(1):77–85.
[38] Statistics Bureau of Urumqi. Urumqi Statistical Yearbook. Urumqi, China: China Statistics Press; 2007.
[39] Statistics Bureau of Urumqi. Urumqi Statistical Yearbook. Urumqi, China: China Statistics Press; 2020.
[40] Zhao Y, Yang D, Zhang X, Xiong C. Spatial distribution characteristics and hot zone patterns of lodging industry in Urumqi. Arid Land Geogr 2016;39(5):1143–52.
[41] Boots B. Developing local measures of spatial association for categorical data. J Geographical Syst. 2003;5(2):139–60.
[42] Dasarathy BV. Nearest neighbor ({NN}) norms:{NN} pattern classification techniques. CA, USA: IEEE CS Press; 1991.
[43] Levine N. Crime mapping and the Crimestat program. Geographical Anal. 2006;38(1):41–56.
[44] Cui C, Wang J, Wu Z, Ni J, Qian T. The socio-spatial distribution of leisure venues: A case study of karaoke bars in Nanjing, China. ISPRS Int J Geo-Inf. 2016;5(9):150.
[45] Batty M. Entropy in spatial aggregation. Geographical Anal. 1976;8(1):1–21.
[46] Liangzhi Y, Wood S. An entropy approach to spatial disaggregation of agricultural production. Agric Syst. 2006;90(1):329–47.
[47] Dixon PM. Ripley's *K* function. Encycl Environ. 2002;3:1796–803.
[48] Hubbard R. A review of selected factors conditioning consumer travel behavior. J Consum Res. 1978;5(1):1–21.
[49] Zhu W. Agent-based simulation and modeling of retail center systems. J Urban Plan Dev. 2015;142(1):04015004.
[50] Getis A, Getis JM. Retail store spatial affinities. Urban Stud. 1968;5(5):317–32.
[51] Gabszwicz JJ. On the nature of competition with differentiated products. Economic J. 1986;96(381):160–72.

[52] Dudey M. A note on consumer search, firm location choice, and welfare. J Ind Econ. 1993;41(3):323–31.
[53] Pashigian BP, Gould ED. Internalizing externalities: The pricing of space in shopping malls. J Law Econ. 1998;41(1):115–42.
[54] Hotelling H. The economics of exhaustible resources. J Political Economy. 1931;39(2):137–75.
[55] Reilly WJ. The law of retail gravitation. New York: Knickerbocker Press; 1931.
[56] Huff David L. Defining and estimating a trade area. J Mark. 1964;28:34–8.
[57] Seim K. An empirical model of firm entry with endogenous product-type choices. Rand J Econ. 2010;37(3): 619–40.
[58] Guimaraes P. Revisiting retail planning policies in countries of restraint of Western Europe. Int J Urban Sci. 2016;20(3):361–80.
[59] Emery J. Bullring: A case study of retail-led urban renewal and its contribution to city centre regeneration. J Retail Leisure Property. 2006;5(2):121–33.
[60] Guy C, Bennison D. Retail planning policy, superstore development and retailer competition. Int J Retail Distrib Manag. 2002;30(9):431–4.
[61] Karamshuk D, Noulas A, Scellato S, Nicosia V, Mascolo C, editors. Geo-spotting: Mining online location-based services for optimal retail store placement. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM; 2013.