

Lin, T. and Xia, J.C. and Ryan, M. and Robinson, T. and Currie, G. and McCarney, G. and Butorac, D. 2018. Exploring passenger rail markets using new station catchment size and shape metrics. *Journal of Spatial Science*. 63 (2): pp. 379-398.

Exploring Passenger Rail Markets Using New Station Catchment Size and Shape Metrics

Ting (Grace) Lin^a, Jianhong (Cecilia) Xia^{*a}, Mark Ryan^a, Todd Robinson^a, Graham Currie^b, Gary MacCarney^a, Donna Butorac^c and Roula Zougheibe^a

^aSchool of Earth and Planetary Sciences, Curtin University, Kent St Bentley, Western Australia

^bPublic Transport Research Group, Institute of Transport Studies, Department of Civil Engineering, Building 60, Monash University, Clayton, Victoria 3800, Australia

^cSchool of Media, Culture and Creative Arts, Curtin University

*Corresponding author: Jianhong(Cecilia) Xia

Phone Number: +61 8+92667563

Email: c.xia@curtin.edu.au

Abstract: *This paper presents a novel spatial market segmentation method to determine key user groups of a train station, such as gender, age and access mode, based on the size and shape of the station's individual catchment area. A user group is considered to be a key market segment for a train station if the size of this group's catchment area is closer to that of the overall catchment area of a train station (i.e., of all user groups combined) and its shape is closer to a circle. Two new indices; area ratio and composite ratio, are developed to quantify the importance of user groups for a train station. This method is applied to identify key user groups at seven train stations in Perth, Western Australia. The study offers a new way to explore the travel behaviours of train users and provides insights for rail transport planning and marketing.*

Keywords: Station catchment size and compactness; spatial market segmentation; area ratio; composite ratio.

1. Introduction

A train station's catchment area refers to the area from which the majority of users will typically be drawn (Dolega et al., 2016). It is vital for understanding latent demand (potential customers) (Banister, 1980), market share (the portion of a market) (Lee and Masao, 1988, p. 17-19) and accessibility (ability to reach) (El-Geneidy and Levinson, 2006). Various catchment area estimation methods have been developed in the literature (Lin et al., 2014, Zhao et al., 2003). The most commonly used method is the proximity method, which uses buffer rings or network-based service areas to determine the spatial extent of a catchment area (Landex and Hansen, 2009). Many studies have been conducted to understand factors affecting catchment size, such as land use diversity and density, transit services and facilities, and accessibility to train stations (García-Palomares et al., 2013, El-Geneidy et al., 2014). However, one area that has not yet been well researched concerns individual differences in the spatial extent of train station catchment areas. This study aims to fill this research gap by comparing the size and shape of train station catchment areas between different user groups, such as age and gender, and developing a novel spatial market segmentation method to determine the spatially dominant user characteristics of train station catchment areas.

Market segmentation is widely used in all sectors of industry. It is an analytical process, driven by customer needs, that dissects the marketplace into submarkets that require different marketing approaches. It has been described as the cornerstone of modern marketing and it is at the heart of marketing strategy, helping to bridge the gap between diverse customer needs and limited business resources (Dibb, 1998). Traditional market segmentation is measured using a defined set of variables: geographic, demographic, psychographic, behavioural (Reid and Bojanic, 2009). However, very little effort has been made towards deriving market segments based on the spatial travel patterns of train users. This study identifies key user groups of train stations based on their size and compactness of catchment area. A key market segment or user group for a train station is one who is willing to travel longer distances and from diverse directions to reach the station. The spatial market segmentation method is a relative measure. This means that if the spread of a key group's catchment area is closer to that of the overall catchment area, (ie. for all users combined), its compactness is closer to one. Two new metrics, the 'area ratio' and the 'composite ratio', were developed to identify the spatially dominant market segments based on four characteristics of users: age, gender, trip directions and access modes.

The paper is structured as follows. Section 2 presents a literature review of catchment area measures, factors affecting the compactness and size of the catchment areas, characteristics of train users and their trips, characteristics of train stations, and the market segmentation methods. Section 3 focuses on discussing our study area, the data collection and analysis methods. The results and findings are presented in Section 4 and discussed in Section 5. The paper ends with conclusions, including suggestions for further study.

2. Research context

2.1. Catchment areas

Catchment area studies have been conducted for many different purposes, such as to identify market share, latent demand estimation, or new store/facility location selection (Goodchild, 1984). Additional studies have focused on gravity model based accessibility analysis (Delamater, 2013, Langford et al., 2012). As a result, a variety of catchment or service area measures for transit stops have been developed since the 1970s (Zhao et al., 2013, Gutiérrez and García-Palomares 2008). They can be categorised into three types. The first is a simple buffer method, which uses a circular buffer with a radius equal to the travel distance or time. An example is the use of rings, (e.g. 400 & 800 meter or 5 & 10 minute radius buffers), to define the walking catchment area of public transport. However, this method has been criticised for not taking geographical surroundings into account and only considering the Euclidean distance, thereby ignoring indirect paths, obstructions, or breaks in the network.

The second type of measure is the road network based service area, which uses travel distances or times similar to the above but along the transport network rather than straight lines (Landex and Hansen, 2009). The network based service area measure avoids the inherent problem of the simple circular buffers, because the shape of the calculated service area corresponds to the road network and surrounding land use. However, it cannot account for the fact that catchment areas can be impacted by a variety of factors, such as the population density of the surrounding area, the location of fare zone boundaries, and even the station facilities. This measure also has the same problem as the simple buffer measure in that it is hard to define the thresholds for generating service areas, (for example, whether to adopt a 10, 15 or 20 minute threshold for access by car).

The third type of measure uses geocoded trip survey data. The travel data are analysed first and then the catchment boundary is delineated based on the origins of the trips. One common method is the convex hull method (Durr et al., 2010), whereby approximately 90% of the survey data are used in order to exclude the effect of outliers (Irvine, 2011, Durr et al., 2010). The trip survey based catchment area definition measure has three steps:

1. geocoding the survey data;
2. excluding the 10% data as outliers
3. delineating the catchment area.

The trip survey collects the origins and destinations of current riders to exactly determine where they come from. The sampling method is critical as a sample of limited size may not reveal the real boundary of the catchment area, but increasing the sample size increases the time and resources involved.

For the buffer and road network based service area measures, it is crucial to define the thresholds. In the literature, the bulk of the catchment area research has been about identifying these thresholds. For example, Zhao et al. (2003) developed distance decay functions to estimate pedestrian walking accessibility to transit stops at the home end and discovered that half a mile, or 800 metres, walk distances were acceptable. This distance has been widely accepted as defining the walk catchment area to transit stations. However, according to El-Geneidy et al. (2014), the 800 metre distance was underestimating the catchment area. Based on their study in the Montreal region, they found 85th percentile walking distances of 1,219 meters and 1,095 meters at the trip origin and destination ends respectively. O'Sullivan and Morrall (1996) identified that the average walking distance to light-rail transit stations in a suburb, (648m with a 75th-percentile distance of 840m), was larger than the distance to stations in the central business district, (326m with a 75th-percentile distance of 419m). Additionally, García-Palomares et al. (2013) found population groups (such as young people and adults, men, immigrants, and public transit captives) who were found to be more likely to walk longer distances and be less sensitive to the effect of distance. Households with high annual incomes (greater than USD 100 K) were willing to travel longer distances compared to households with lower incomes (Hochmair, 2015).

2.2. Market segmentation

Since it was first introduced by Smith (1956), market segmentation has become a central concept in both marketing theory and practice. Although many definitions of market segmentation have been proposed since 1956, the original definition by Smith is still agreed upon and adopted by most researchers. It is: “*Market segmentation involves viewing a heterogeneous market as a number of smaller homogeneous markets, in response to differing preferences, attributable to the desire of consumers for more precise satisfaction of their varying wants*” (Smith, 1956, Wedel and Kamakura, 2012). It is the process of splitting customers/potential customers within a market into different groups/segments with similar needs (McDonald, 2012). Through dissecting the marketplace into submarkets, market segmentation allows organizations to focus their resources more effectively and with a greater chance of success. It can help in product and service development and marketing. Weinstein (2013) stated it as “*the key to marketing success*” and “*segmentation imperative*”.

Segmentation is essentially a grouping task and there are a large number of methods available to undertake this grouping. Wedel and Kamakura (2012) classified the current segmentation methods and techniques into four categories which are combinations of two major classifications, (a-priori or post-hoc and descriptive or predictive). The resulting categories are: a-priori descriptive, post-hoc descriptive, a-priori predictive and post-hoc predictive. In a-priori descriptive segmentation, the type and number of segments are determined before data collection; whilst in post-hoc descriptive methods, the segments are identified after data collection based on grouping heterogeneous data

together, (clustering analysis being the most popular approach). In a-priori predictive segmentation, the definition of segments is required based on a set of criteria and then subsequently, the predictive models are used to describe the relationship between the segment membership and a set of independent variables. In post-hoc predictive segmentation, the identification of the segments is on the basis of the estimated relationship between a dependent variable and a set of predictors. AID (Automatic Interaction Detection), CART (Classification and Regression Trees), clusterwise regression and Artificial Neural Network (ANN) are all methods that can be used in the post-hoc predictive segmentation process. However, these approaches result in segmentations that are primarily from an empirical, statistical or mathematical perspective. The literature review has indicated a lack of research that considers the segmentation problem from a spatial perspective. This paper, to the knowledge of the authors, will be the first attempt to fill this gap.

3. Methods

The spatial market segmentation analysis in this paper was undertaken in two steps. The first step was to delineate and measure the size and shape of catchment areas for a range of different user groups, based on age (eg. young or elderly), sex (male or female), direction of travel (inbound or outbound), and station access mode (eg. bus or car). In the second step, two new indicators were developed, the area ratio and the composite ratio, to determine which segments were, spatially, the key market segments. The methodology is explained further below.

3.1. Study area and data collection methods

The spatial segmentation methodology was applied to seven train stations within the Perth metropolitan area, Western Australia (see Figure 1). Perth is the state capital of Western Australia and has five train lines - Armadale, Fremantle, Joondalup, Midland and Mandurah - and one spur line, the Thornlie line, which is off the Armadale line. Perth has 70 train stations on 173 kilometres of track (Australian Department of Infrastructure and Transport, 2014). The seven train stations, (Cannington, Claremont, Greenwood, Midland, Murdoch, Warnbro and Warwick stations) were selected in consultation with industry partners; the Department of Transport, (DoT), the Department of Planning, (DoP) and the Public Transport Authority, (PTA). These stations were identified by the industry partners as having unique features or problems that required further research.



Figure 1: The study area

Intercept surveys were conducted at the above seven stations to collect train user trip information and to determine their level of satisfaction with train services and facilities. The surveys were undertaken on 31 July 2012 and 1 August 2012 between 6:00AM and 4:00PM and 19 and 20 September 2013 between 7:00AM and 12:30PM. People were interviewed while they were waiting to board and a total of 1269 responses were collected. Table 1 provides a breakdown of the survey responses for each station by segmentation variable, (ie. age, gender, trip direction and travel mode). Generally, the elderly group (over 60) sample size is smaller than the middle-aged (25-59) and young (18-24) group sample sizes. Except for the Greenwood station survey, which is dominated by middle-aged users, there is no significant difference between the young and middle-aged survey sizes at the other stations. The inbound (towards city centre) survey sample size is much larger than the outbound (away from city centre) sample due to the majority of train trips being to the central area and the seven stations being outside this central area. Claremont and Greenwood stations tend to be more likely to accommodate park and ride (PnR) users, whilst at Warwick and Murdoch stations, Bus and Ride (BnR) is the dominant travel mode.

Table 1 Survey sample characteristics by train user types and stations

Station	The sample size by train user and station						
	Cannington	Claremont	Greenwood	Midland	Murdoch	Warnbro	Warwick
All*	152	107	161	120	117	159	135
Young (18-24)	61	40	37	45	57	83	64
Middle (25-59)	67	48	106	40	37	50	53
Elderly (over 60)	14	14	14	29	20	17	15
Male	91	51	79	56	64	78	58

Female	80	56	82	64	53	81	77
Inbound (towards city centre)	132	82	152	120	97	153	116
Outbound (away from city centre)	20	25	9	0	20	6	19
Bus and Ride	57	8	3	24	52	48	43
Park and Ride	40	42	79	56	22	49	42

*The number of sub categories don't sum to the total number due to missing information

3.2. Delineation of train station catchment areas

The research aims to better understanding not only the size of a catchment area but also its shape. Hence the determination of spatial boundaries of catchment areas is crucial. In this paper, the catchment areas were determined by initially plotting the survey data, ie. by geocoding the location of each trip origin using Google application programming interface (API). Then, the minimum bounding geometry, or Convex Hull approach (Cervero et al. 1995b, Guerra, Cervero, and Tischler 2011), was adopted to determine the boundaries of the catchment areas. This method uses computational geometry theory to derive a convex hull containing 90 percent of the trip origin location points, (as 10 percent of the sample data are considered as spatial outliers and are removed). The boundary of the catchment area is a series of line segments joining the outermost points so that all remaining points are enclosed (De Berg et al., 2000). Figure 2 illustrates the convex hull method. An advantage of this method is that it captures the spatial boundary of the catchment area based on the location of individual trip origin data in a disaggregated manner, thereby reflecting the actual shape of the catchment more exactly. The disadvantage of this method is that it needs a relatively large sample size in order to make a catchment area representative and it is sensitive to spatial outliers. In order to determine the key spatial segments, a separate catchment area was determined for each segmentation variable (ie. age, gender, trip direction and travel mode) for each station. The size of each of these separate train station catchment areas was then determined using ArcGIS 10.2 software by ESRI (Esri, 2010).

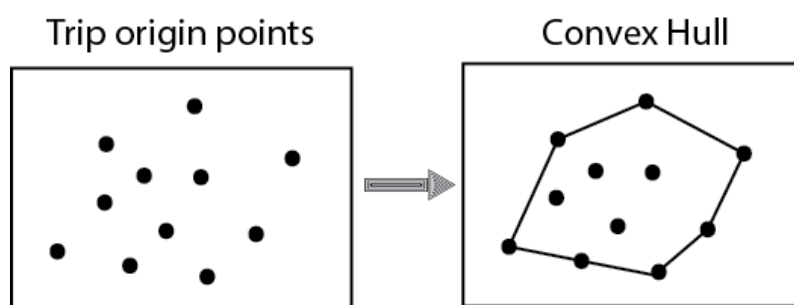


Figure 2: An illustration of the minimum bounding geometry approach

3.3. Catchment shape measures: compactness measure

Shape measures quantify aspects of the catchment area shape, such as compactness of shape (Li et al., 2013) and can be used to exhibit transportation efficiency and homogeneity of the regions around a station in terms of symmetry of network and accessibility of service (Maceachren, 1985). Maceachren (1985) stated that shape measures allow “*a measure of shape uniqueness by which any shape can be distinguished from all other shapes and similar shapes result in similar descriptions.*” Shape measures have widespread applications especially in landscape ecology and geography (Taylor, 1973, BÉLanger and Eagles, 2001, Gardoll et al., 2000). Hundreds of metrics exist for measuring the characteristics of shapes (Angel et al., 2010).

Compactness is acknowledged as one of the most intriguing and important properties of a shape (Angel et al., 2010) and is widely used as a descriptor in various disciplines (Li et al., 2013). It quantifies how compact a shape is. Maceachren (1985) categorised the compactness of geographic shapes into four groups: perimeter-area measurements, single parameters of related circles, direct comparison to a standard shape and dispersion of elements of an area around a central point.

This paper adopted the compactness measurement of single parameters of related circles because a station catchment area has no hole inside and is scale-invariant. A train station is easier to travel to if the shape of its catchment area is roughly circular rather than long and thin. The compactness of a shape measure is defined as (Cole, 1964):

$$C = A / \pi R^2 \quad (1)$$

where C is the compactness of a catchment area; A is the area of a catchment area; and R is the radius of the smallest circle that encloses the catchment area (Figure 3). The compactness of a catchment area of a train station is between 0 and 1. One means the catchment area is a perfect circle, ie. the catchment is “completely isotropic”. If the compactness value is close to 0, the catchment area is almost a line, which means people come from only one orientation, (eg north/south or east/west) to reach a train station and the catchment is “completely anisotropic”. A station is not necessarily located at the centre of its catchment area. The above compactness measure did take this into consideration. The compactness was also calculated for each segmentation variable (ie. age, gender, trip direction and travel mode) for each station. The results of the compactness analysis for the seven stations in the case study are presented in Section 4.

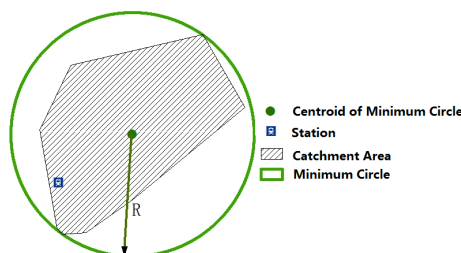


Figure 3: An illustration of compactness calculation

3.4. The spatial market segmentation concept

The fundamental assumption of the spatial market segmentation method adopted in this paper is that the size and compactness of the station catchment area varies depending on the user group and that the bigger the size and the higher the compactness of the catchment area of a particular user group, the more attractive the station is for that group. This means that if a particular user group, say young train users, is willing to travel longer distances and from various directions to reach a transit station, the station is more attractive to this user group than to other user groups.

The key research question is therefore: how big or how compact does a particular user group's station catchment area need to be for that user group to be considered a key market segment, ie. how close to the overall catchment size? In order to identify dominant characteristics of a station catchment area, two new measures were developed: a disaggregate market segment area ratio and a composite ratio.

- *Area ratio*

The area ratio is the ratio of the size of the catchment area for a particular subgroup or market segment, such as Female, to the size of the overall catchment area for all subgroups combined.

$$r_{Aij} = A_{subij} / A_{wholej} \quad (4)$$

where r_{Aij} is the area ratio of subgroup i at station j ; A_{subij} is the catchment area for subgroup i at station j ; A_{wholej} is the catchment area for all subgroups combined at station j ;

The area ratio is between 0 and 1. The higher the area ratio value, the more dominant a market segment is in terms of catchment area. For example, the area ratio of middle-aged is 0.98 for one of the train stations, which is almost the size of the overall catchment area. This could mean that the middle-aged travellers are the dominant user group at that station and are willing to travel longer distances to reach the station. However, this measure has the disadvantage of only indicating travel distances and not direction of travel. Due to this limitation, another measure was developed, the composite ratio, which considers both distance and direction.

- *Composite ratio*

The composite ratio is calculated as follows:

$$r_{Comij} = r_{Aij} \times C_{ij} \quad (5)$$

where r_{Comij} is the composite ratio of subgroup i at station j ; r_{Aij} is the area ratio for subgroup i at station j ; C_{ij} is the compactness of a catchment area for subgroup i at station j ;

The composite ratio is between 0 and 1. The higher the composite ratio value, the more dominant a market segment is in terms of the size and compactness of catchment area. Therefore, the composite ratio reflects both trip distance and direction. A dominant or key market segment is more likely to travel longer distance and from diverse directions.

4. Results

4.1. The size and compactness of a catchment area

The size and compactness of the catchment areas were calculated for all users combined and separately for each user group at each of the seven stations (See Table 2 and Figure 4). Midland station has the largest overall catchment area, nearly 23 times larger than the smallest one: the catchment area of Claremont station. This is probably because Midland station is located at the end of Midland rail line and serves a large urban and semi-urban area. On the other hand, Claremont station serves a small area constrained by the Swan River, is a short distance to adjacent train stations (0.7km to the nearest inbound direction station, Showgrounds, and 1.1km to the nearest outbound direction station, Swanbourne and with a large part of the catchment area covered by other types of land use, such as lakes, park lands and recreation facilities, instead of residential areas (Shao et al., 2015). The catchment area of Cannington station is the second largest. This could be due to the presence of the Westfield Carousel shopping centre, which is the largest shopping centre in Western Australia and located approximately 600m from the train station. When comparing the catchment areas of the three age groups, the catchment areas of the elderly are, with the exception of Greenwood, the smallest. The young user group catchment areas are the largest at five stations.

Table 2 Area size and compactness of the seven station catchment area

Station	The size of a catchment area (Km ²)										
	Cannington	Claremont	Greenwood	Midland	Murdoch	Warnbro	Warwick	Including Midland		Excluding Midland	
								Mean	SD	Mean	SD
All	165.06	46.56	80.85	1062.35	112.05	74.34	81.95	231.88	368.09	93.47	40.81
Young	136.01	37.13	44.47	584.82	86	68.68	80.97	148.3	195.19	75.54	35.45
Middle	104.93	45.88	36.92	739.44	78.82	46.63	65.45	159.72	256.69	63.11	25.52
Elderly	69.55	21.2	43.15	337.92	52.57	38.81	23.21	83.77	113.3	41.42	18.25
Male	139.04	34.45	51.6	662.03	101.37	60.11	80.84	161.35	223.48	77.9	37.93
Female	113.36	46.49	51.26	800.48	90.92	61.47	69.45	176.2	276.27	72.16	25.56
Inbound	155.59	45.95	57.46	1062.35	107.33	74.34	65.33	224.05	371.51	84.33	40.65
Outbound	78.2	21.73	29.93	0	48.11	10.61	56.47	35.01	27.45	40.84	24.86
BnR (Bus and Ride)	104.97	23.97	0	323.17	73.02	41.26	63.06	89.92	108.35	51.05	37.34
PnR (Park and Ride)	125.59	41.95	60.9	693.11	47.48	61.81	81.95	158.97	237.18	69.95	30.58
Mean (Subgroups)	114.14	35.42	41.74	578.15	76.18	51.52	65.19				

SD (Subgroups)	28.28	10.67	18.43	313.68	22.65	19.55	18.15				
Station	The compactness of a catchment area										
	Cannington	Claremont	Greenwood	Midland	Murdoch	Warnbro	Warwick	Including Midland		Excluding Midland	
								Mean	SD	Mean	SD
All	0.51	0.28	0.52	0.53	0.56	0.33	0.69	0.49	0.14	0.48	0.15
Young	0.59	0.22	0.49	0.34	0.54	0.3	0.68	0.45	0.17	0.47	0.18
Middle	0.41	0.28	0.37	0.44	0.44	0.44	0.58	0.42	0.09	0.42	0.1
Elderly	0.54	0.26	0.42	0.27	0.4	0.41	0.49	0.4	0.1	0.42	0.1
Male	0.52	0.23	0.55	0.33	0.55	0.31	0.68	0.45	0.16	0.47	0.17
Female	0.49	0.28	0.5	0.46	0.59	0.29	0.62	0.46	0.13	0.46	0.15
Inbound	0.57	0.28	0.5	0.53	0.54	0.33	0.58	0.48	0.12	0.47	0.13
Outbound	0.5	0.31	0.44		0.38	0.34	0.48	0.41	0.08	0.41	0.08
BnR	0.55	0.16		0.57	0.53	0.4	0.56	0.46	0.16	0.44	0.17
PhR	0.56	0.28	0.53	0.38	0.52	0.33	0.69	0.47	0.15	0.49	0.15
Mean (Subgroups)	0.53	0.26	0.48	0.42	0.50	0.35	0.60				
SD	0.04	0.05	0.06	0.10	0.07	0.05	0.08				

Station	Station boarding September 2011						
	Cannington	Claremont	Greenwood	Midland	Murdoch	Warnbro	Warwick
Total Daily Station							
Boarding*	3,165	1,967	2,143	3,899	7,898	2,735	5,867
Bus to Train transfers	1,129	227	4	1,211	4,430	736	2,480

Regarding compactness (See Table 2), Claremont station's catchment area has the longest and thinnest shape with the lowest compactness value of 0.28. It also had the lowest number of boarders. Warwick station had the highest compactness value (0.69) and the second largest number of boarders. It could be that compactness has some positive relationship with train station ridership. Furthermore, the compactness value for the overall catchment area is not always the highest when compared to the value for an individual user group. For example, at Cannington station, the catchment area for young (18-25 years old) users has the largest compactness value (0.57), which means that young train users are more likely to travel from diverse directions to reach the station than the other age groups. However, the compactness of the young adult group catchment areas (SD = 0.17) was found to vary relatively higher among the seven stations compared to other age groups. Midland station has the highest variation between different user groups (SD = 0.10) compared with Cannington with the lowest variation (SD = 0.04). This is illustrated on Figure 4. The compactness of a catchment area can also show station accessibility problems for certain user groups. For example, although Warwick station generally has good accessibility by most user groups, the elderly mostly access the station from a narrow area to the southwest. This could be due to less demand or some barriers hindering

elderly access from the other directions (See Figure 4).

Figure 4 provides an illustration of the spatial boundary of the seven stations and their subgroup catchment areas. The size and shape of catchment areas vary greatly among train stations and their subgroups. For example, Greenwood station has the highest variation of the size and shape of subgroup catchment areas compared to the overall catchment area. The spatial boundary of BnR catchment area wasn't displayed in the graph due to lack of sufficient data. Elderly and female train users tended to come from areas north of the station while young and male train users were more likely to come from areas south of the station. However, the catchment area of the middle-aged group was distributed in an east-west direction. Generally, young people had relatively larger catchment areas than elderly and middle-aged users. The size and compactness of PnR and BnR user group catchment areas were heavily influenced by the surrounding road network and bus routes as well as the frequencies of bus and train services. For example, at Greenwood, the BnR catchment area is much smaller than the PnR catchment area due to the station only being serviced by one bus route. The inbound and outbound catchment areas were not necessarily directly related to their direction of travel. In some instances, people chose stations that were located further away from their destination.

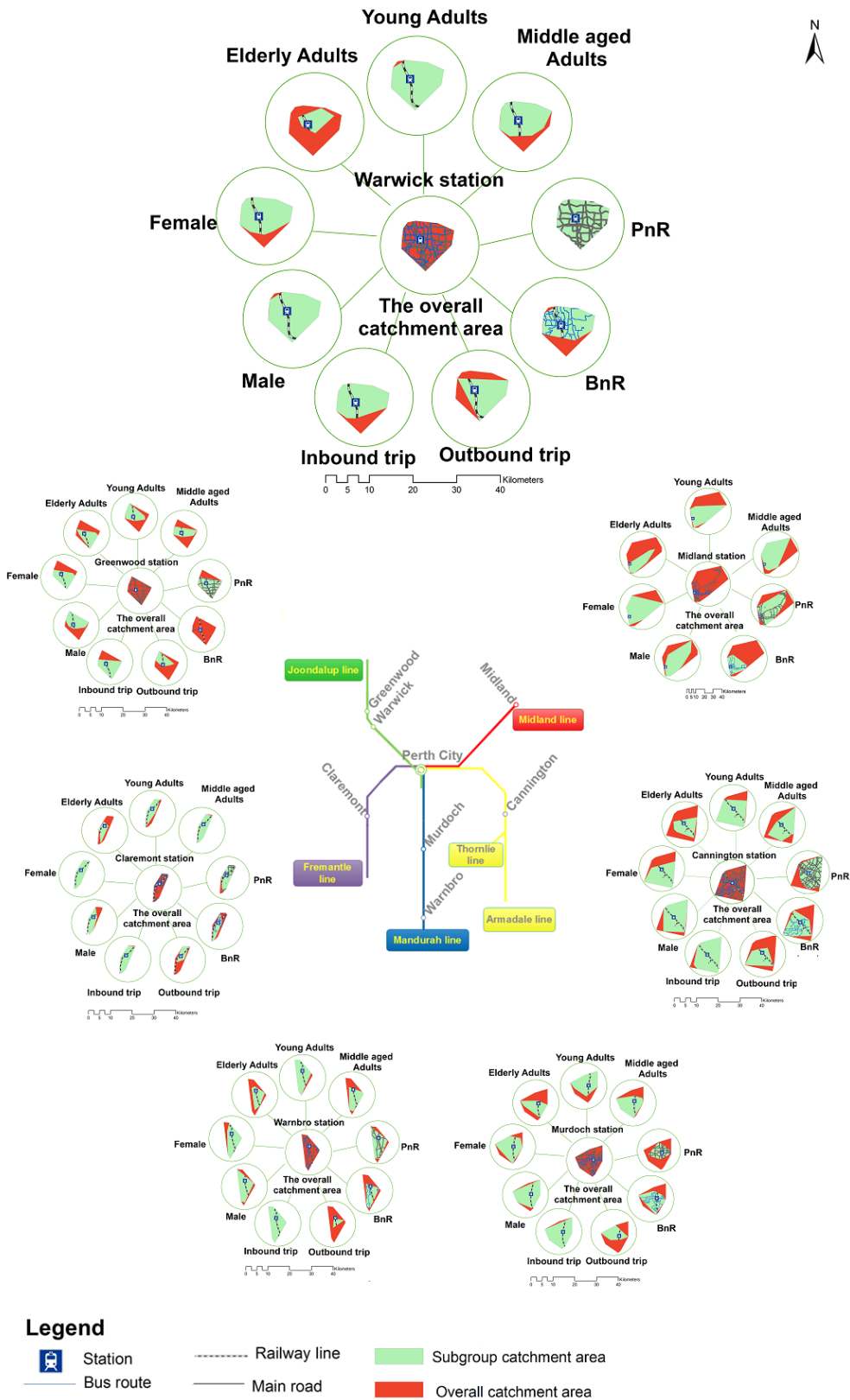


Figure 4: The spatial boundary of the catchment area of train stations by subgroups (the size of the icons representative of size of catchments)

4.2. Disaggregate market segment area ratio and composite ratio

As discussed in Section 3.5, area ratio and composite ratio measures were developed as part of the spatial segmentation analysis. The **area ratio** is the ratio of the size of a subgroup catchment to the size of the overall catchment. The purpose of this ratio is to identify a key market segment, ie. one that is willing to travel a longer distance to reach a train station. The area ratio ranges between 0 and 1, the closer to 1, the closer the size of that user group catchment area is to the size of the overall catchment area and the more likely it is that user group is a key market segment (See Figure 4 and Table 5). Table 5 presents the area ratio of the user group catchment area for the seven train stations. A number of useful insights emerge from this analysis:

- The area ratio of the female user group for Claremont station is *one*. This means the size of the female catchment area is equal to the size of the overall catchment area. In other words, the female catchment area spatially defines the overall station catchment area. Therefore, we might reasonably conclude that the female group could be considered as one of the key market segments for Claremont station.
- The inbound, middle aged and PnR user groups are also relatively important market segments for Claremont station.
- Greenwood station has the lowest area ratio of its subgroup catchment area compared to other train stations whilst Warwick station has the highest. Nevertheless, the PnR user group was found to be a key market segment for both stations.
- Most stations have a larger PnR than BnR area ratio except for Murdoch station.
- It can also be seen in Figure 4 that the PnR catchment area is much smaller than BnR catchment area for Murdoch station.

The **composite ratio** is the area ratio weighted by the compactness of the catchment area. Again, the composite ratio ranges between 0 and 1, the higher the value, the more likely that user group is to be a key market segment of a transit station. The results for the seven stations are also presented in Table 5. The key market segments determined based on the composite ratio vary slightly from the ones determined based on the area ratio. For example, compared to the area ratio results, the composite ratios indicate that the young-age group becomes more important than the male group for Cannington station, which means that the shape of the young-aged catchment area is more compact than the shape of male user group catchment area, (the most dominant area from the area ratio results). Although while in general the two methods provide similar results with respect to dominant user groups, the composite ratio is more sensitive to the shape of catchment area, for example if a train station has a catchment area with a long narrow shape, such as Claremont station. The size and compactness of subgroup catchment areas, such as, female, middle and inbound at Claremont station are very similar to each other, hence the composite ratios of those are similar to each other (See Figure 4 and Table 2&3).

Table 3 Spatially dominant market segments for each station

Rank	Area ratio						
	Cannington	Claremont	Greenwood	Midland*	Murdoch	Warnbro	Warwick
1	InBound (0.94)	Female (1.00)	PnR (0.75)	InBound (1.00)	InBound (0.96)	InBound (1.00)	PnR (1.00)
2	Male (0.84)	InBound (0.99)	InBound (0.71)	Female (0.76)	Male (0.90)	Young (0.92)	Young (0.99)
3	Young (0.82)	Middle (0.99)	Male (0.64)	Middle (0.70)	Female (0.81)	PnR (0.83)	Male (0.98)
4	PnR (0.76)	PnR (0.90)	Female (0.63)	PnR (0.65)	Young (0.77)	Female (0.83)	Female (0.85)
5	Female (0.69)	Young (0.80)	Young (0.55)	Male (0.62)	Middle (0.70)	Male (0.81)	Middle (0.80)
6	BnR (0.64)	Male (0.74)	Elderly (0.53)	Young (0.55)	BnR (0.65)	Middle (0.63)	InBound (0.80)
7	Middle (0.64)	BnR (0.51)	Middle (0.46)	Elderly (0.31)	Elderly (0.47)	BnR (0.56)	BnR (0.77)
8	OutBound (0.47)	OutBound (0.47)	OutBound (0.37)	BnR (0.30)	OutBound (0.43)	Elderly (0.52)	OutBound (0.69)
9	Elderly (0.42)	Elderly (0.46)			PnR (0.42)	OutBound (0.14)	Elderly (0.28)

Rank	Composite ratio						
	Cannington	Claremont	Greenwood	Midland	Murdoch	Warnbro	Warwick
1	InBound (0.54)	Female (0.28)	PnR (0.40)	InBound (0.53)	InBound (0.52)	InBound (0.33)	PnR (0.68)
2	Young (0.48)	Middle (0.28)	InBound (0.36)	Female (0.35)	Male (0.50)	Middle (0.28)	Young (0.67)
3	Male (0.44)	InBound (0.28)	Male (0.35)	Middle (0.31)	Female (0.48)	Young (0.28)	Male (0.67)
4	PnR (0.42)	PnR (0.25)	Female (0.32)	PnR (0.25)	Middle (0.42)	PnR (0.28)	Female (0.53)
5	BnR (0.35)	Young (0.18)	Young (0.27)	Male (0.20)	BnR (0.34)	Male (0.25)	Middle (0.47)
6	Female (0.33)	Male (0.17)	Elderly (0.22)	Young (0.19)	Middle (0.31)	Female (0.24)	InBound (0.47)
7	Middle (0.26)	OutBound (0.15)	Middle (0.17)	BnR (0.17)	PnR (0.22)	BnR (0.22)	BnR (0.43)

8	OutBound (0.24)	Elderly (0.12)	OutBound (0.16)	Elderly (0.08)	Elderly (0.19)	Elderly (0.21)	OutBound (0.33)
9	Elderly (0.23)	BnR (0.08)			OutBound (0.17)	OutBound (0.05)	Elderly (0.14)

5. Discussions

5.1. Spatial market segmentation

Traditional market segmentation methods derive the dominant characteristics of users based on the frequency or probability of those characteristics occurring in the data (Witten and Frank, 2005). However, the spatial market segmentation methods developed in this study identify key market segments by exploring the size and compactness of their train station catchment areas. The area ratio and composite ratios were developed to identify the dominant characteristics of train station catchment areas from a spatial perspective, (distance and direction). The area ratio identifies which user groups travelled longer distances to access stations. The composite ratio identifies user groups that travelled both longer distances and from diverse directions. User groups who are willing to travel longer distances to reach a station do not necessarily access the train station from diverse directions. Identifying the spatially dominant characteristics of users could assist transport planners and operators in managing the demand and supply sides of train services. For example, on the demand side, information such as the core users (dominant characteristics) could be a valuable source for targeted marketing to further promote train services or to market some relevant business by using customised business advertisements at or around train stations. On the supply side, information about the catchment areas of non-dominant users could facilitate an understanding of the problems or barriers to train service use by certain user groups. Therefore, further interventions or strategies could be put in place to encourage these users to use train services (Mulley et al., 2012).

5.2. Characteristics of a catchment area

The most noticeable result from this study was that different user subgroups have catchment areas of varying size and shape, which may influence or reflect their travel behaviour. For example, the size of the middle-aged catchment area is larger than that of other age groups. However, the compactness of the young user group catchment is higher than that of other age groups. The middle-aged group travel longer distances to reach a train station, while the young user group travel from more diverse directions to access a train station compared to other age groups. According to the ABS (2013) social trends survey, young people had one of the highest shares of travel by public transport to get to work or study (28%), whilst the middle-aged population (55-64 years) were the most likely to drive to work or study (78%). Delbosch and Currie (2014) also identified reasons why young Australians may be turning their back on the car, such as a reduction in those getting a driving license, a change in the social status of the car, greater awareness of the environment and the role of electronic

communications. All these may explain why the young user group catchment area is more compact than that of older age groups. Although PnR catchment sizes are larger than those for BnR, the compactness of the PnR and BnR catchment areas is similar. This may indicate that PnR is more flexible than BnR, ie. it allows PnR users to drive longer distances to access a train station.

Good train services alone may not attract diverse and intense usage. An integrated transport system has been proven to be a more efficient way to increase the mobility of a community (Cervero and Kockelman, 1997, Ewing and Cervero, 2010). The spatial market segmentation methodology in this paper used a limited number of user characteristics, such as age and gender, selected based on the data available to this study. The key segments identified were therefore limited to these pre-selected segments. The methodology described can however be used to identify other potential key market segments, such as disaggregation by income level, ethnicity and affordability or social mix which are all known to affect rail ridership (Lucas and Jones, 2012). In addition, Walk and Ride (WnR) was excluded from the analysis due to limitations in the survey. This is clearly another area for future analysis.

5.3. Catchment area measures

There has been some debate over the validity and limitations of the convex-hull method. Some studies suggest that convex hull polygons are inferior in the applications of potential path areas and activity spaces. The main reasons are 1) the convex activity-space polygon might overestimate or underestimate the activity space due to sampling limitation (Chaix et al., 2012). and 2) the convex hull polygon method is sensitive to spatial outliers (Thériault et al., 1999). However, the choice of methods depends on the purpose of the study. In this research, the catchment area of a train station was defined as an area within which local residents could **potentially** use train services. Therefore, it would be reasonable to include some areas where residents haven't used train services yet but could potentially use the services in the future. In addition, in order to avoid spatial outliers, 10% of data, which could be spatial outliers, were removed for the analysis. Other methods, such as spatial outlier detection algorithms could be used to improve the validity of the convex hull polygon methods (Lu et al., 2003, Shekhar et al., 2003). In the future, a study will be developed to systematically evaluate the efficiency of spatial outlier detection methods on the improvement of the convex hull polygon method. The compactness is a good measure of the shape of catchment areas, which can indicate the trip direction. However, this direction refers to the centre of the catchment area rather than the train station itself, (which is rarely in the centre of its catchment area).

The size and shape of a catchment area are sensitive to the sample size because its spatial boundary is determined based on the origin locations of individual travellers. Although the survey was conducted from 6:00am to 4:00pm and covered the period when most trips used a station as the origin station, the

sample size of some of the subgroups, (eg. the elderly), is not that large. Therefore, the catchment areas may not be fully representative. Further analysis has been conducted to test the sample size influence on the catchment area by comparing the elderly group catchment areas derived from 2012 data only with those derived from the 2012 plus 2013 data, for all data and for the outliers removed. The results, presented in Table 6, show that both the size and shape values change with changes in the sample size. For all stations except Cannington station, the catchment area size increase as the sample size increases. There is a notable size change at Midland, Claremont and Greenwood stations. Interestingly, Midland has a very large catchment area, whilst Claremont station has a very small catchment area. Maybe when the catchment area is extremely large or small, it is more sensitive to spatial outliers. For Greenwood station, the threefold increase of the sample size (5 to 14) might be the reason for the change of the catchment area size. In addition, the size and shape of catchment areas may be closely related to residential locations. For example, trip directions mostly align with residential locations. Figure 6 illustrates the residential location of the elderly. Many retired people live on the southeast side of Midland station. The shape of the elderly catchment area of Midland station was elongated in this direction (See Figures 5&6).

Table 6 Elderly catchment area comparisons

The elderly segmentation										
Station	2012 only				2012+2013				Change (2012/(2012+2013))	
	Sample Size(All data)	Sample Size(Outliers removed)	Size (Km2)	Shape	Sample Size(offal data)	Sample Size(Outliers removed)	Size (Km2)	Shape	Size Change	Shape Change
Cannington	16	13	94.35	0.45	20	14	69.55	0.54	1.36	0.83
Claremont	11	10	7.28	0.47	16	14	21.2	0.26	0.34	1.8
Greenwood	5	5	10.35	0.26	18	14	43.15	0.42	0.24	0.62
Midland	28	21	85.68	0.56	33	29	337.92	0.27	0.25	2.07
Murdoch	15	15	48.78	0.37	20	20	52.57	0.4	0.93	0.93
Warnbro	13	12	34.43	0.4	18	17	38.81	0.41	0.89	0.99
Warwick	14	13	18.91	0.61	18	15	23.21	0.49	0.81	1.26

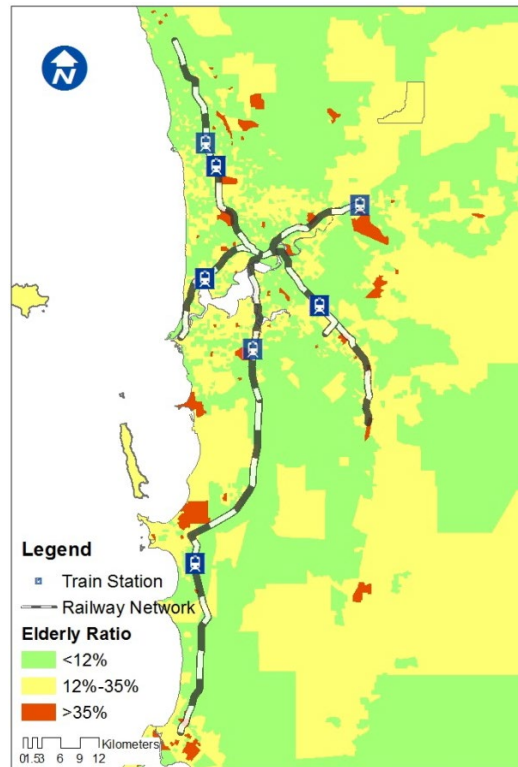


Figure 6. The map of the elderly distribution

6. Conclusions

Identifying the size and compactness of train station catchment areas is an important input to the understanding distance travelled and the trip direction of transit users (El-Geneidy et al., 2014). It is also important to investigate the different types of transit users and the distances and directions they travel for marketing purposes. In this paper GIS techniques have been applied to derive and visualise the spatial distribution of catchment areas by various user groups, which have been shown to clearly relate to distance travelled and travel direction. If train users accessed a train station from a small, thin and long catchment area, this might indicate that this station could have some potential accessibility constraints or problems hindering users. If the size and shape of the catchment areas for different train user groups vary significantly or a train station only attracts a certain group of users from a certain direction, it may indicate that some characteristics of a station or its catchment area are playing a negative role in attracting this group of train users. In addition, this study has developed a novel method for identifying the spatially dominant market segments of train services. Understanding the size and compactness of catchment areas by different user groups and target markets of stations is a vital component of integrated transport planning and promoting train services.

Reference

- ABS. 2013. *Car Nation* [Online]. Available: <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features40July+2013> [Accessed August 7th 2016].
- ANGEL, S., PARENT, J. & CIVCO, D. L. 2010. Ten compactness properties of circles: measuring shape in geography. *Canadian Geographer / Le Géographe canadien*, 54, 441-461.
- AUSTRALIAN DEPARTMENT OF INFRASTRUCTURE AND TRANSPORT 2014. Urban public transport: updated trends.
- BANISTER, D. 1980. *Transport mobility and deprivation in inter-urban areas / David Banister*, Farnborough, Eng., Farnborough, Eng. : Saxon House.
- BÉLANGER, P. & EAGLES, M. 2001. The compactness of federal electoral districts in Canada in the 1980s and 1990s: an exploratory analysis. *Canadian Geographer / Le Géographe canadien*, 45, 450-460.
- CERVERO, R. & KOCKELMAN, K. 1997. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2, 199-219.
- CHAIX, B., KESTENS, Y., PERCHOUX, C., KARUSISI, N., MERLO, J. & LABADI, K. 2012. An interactive mapping tool to assess individual mobility patterns in neighborhood studies. *American journal of preventive medicine*, 43, 440-450.
- COLE, J. P. 1964. Study of major and minor civil divisions in political geography. *The 20th International Geographical Congress*. London.
- DE BERG, M., VAN KREVELD, M., OVERMARS, M. & SCHWARZKOPF, O. C. 2000. *Computational geometry*, Springer.
- DELAMATER, P. L. 2013. Spatial accessibility in suboptimally configured health care systems: A modified two-step floating catchment area (M2SFCA) metric. *Health & Place*, 24, 30-43.
- DELBOSC, A. & CURRIE, G. 2014. Using discussion forums to explore attitudes toward cars and licensing among young Australians. *Transport Policy*, 31, 27-34.
- DIBB, S. 1998. Market segmentation: strategies for success. *Marketing Intelligence & Planning*, 16, 394-406.
- DOLEGA, L., PAVLIS, M. & SINGLETON, A. 2016. Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services*, 28, 78-90.
- DURR, P., GRAHAM, K. & EADY, S. 2010. GIS mapping of cattle market service areas using the National Livestock Identification System. Australian Animal Health Laboratory, Geelong, Australia
- CSIRO Livestock Industries, Armidale, Australia.
- EL-GENEIDY, A., GRIMSRUD, M., WASFI, R., TÉTREAULT, P. & SURPRENANT-LEGAULT, J. 2014. New evidence on walking distances to transit stops: identifying redundancies and gaps using variable service areas. *Transportation*, 41, 193-210.
- EL-GENEIDY, A. & LEVINSON, D. 2006. Access to destinations: Development of accessibility measures.
- ESRI. 2010. *World street map*. Esri.
- EWING, R. & CERVERO, R. 2010. Travel and the built environment. *Journal of the American planning association*, 76, 265-294.
- GARCÍA-PALOMARES, J. C., GUTIÉRREZ, J. & CARDOZO, O. D. 2013. Walking accessibility to public transport: an analysis based on microdata and GIS. *Environment and Planning B: Planning and Design*, 40, 1087-1102.
- GARDOLL, S. J., GROVES, D. I., KNOX-ROBINSON, C. M., YUN, G. Y. & ELLIOTT, N. 2000. Developing the tools for geological shape analysis, with regional- to local-scale examples from the Kalgoorlie Terrane of Western Australia. *Australian Journal of Earth Sciences*, 47, 943-953.
- GOODCHILD, M. F. 1984. Ilacs: A location allocation model for retail site selection. *Journal of Retailing*, 60, 84-100.
- GUTIÉRREZ, J. & GARCÍA-PALOMARES, J. C. 2008. Distance-measure impacts on the calculation of transport service areas using GIS. *Environment and Planning B: Planning and Design*, 35, 480-503.

- HOCHMAIR, H. H. 2015. Assessment of Bicycle Service Areas around Transit Stations. *International Journal of Sustainable Transportation*, 9, 15-29.
- IRVINE. 2011. *Analyzing the Market Share of Commuter Rail Stations using LEHD Data* [Online]. TRB. Available: http://onlinepubs.trb.org/onlinepubs/Conferences/2011/Census/Presentations/10-26_830-10amCervenka/2Eichler.pdf [Accessed 10/3/1014 2014].
- LANDEX, A. & HANSEN, S. 2009. GIS-based Approaches to Catchment Area Analyses of Mass Transit. *ESRI International User Conference*.
- LANGFORD, M., HIGGS, G. & FRY, R. 2012. Using floating catchment analysis (FCA) techniques to examine intra-urban variations in accessibility to public transport opportunities: the example of Cardiff, Wales. *Journal of Transport Geography*, 25, 1-14.
- LEE, G. C. & MASAO, N. 1988. *Market share analysis: Evaluating competitive marketing effectiveness*, Kluwer Academic Publishers.
- LI, W., GOODCHILD, M. F. & CHURCH, R. 2013. An efficient measure of compactness for two-dimensional shapes and its application in regionalization problems. *International Journal of Geographical Information Science*, 27, 1227-1250.
- LIN, T., PALMER, R., XIA, J. & MCMEEKIN, D. A. Automatic generation of station catchment areas: A comparison of Euclidean distance transform algorithm and location-allocation methods. 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2014, 2014. 963-967.
- LU, C.-T., CHEN, D. & KOU, Y. Algorithms for spatial outlier detection. *Data Mining, 2003. ICDM 2003. Third IEEE International Conference on, 2003. IEEE*, 597-600.
- LUCAS, K. & JONES, P. 2012. Social impacts and equity issues in transport: an introduction. *Journal of Transport Geography*, 21, 1-3.
- MACEACHREN, A. M. 1985. Compactness of Geographic Shape: Comparison and Evaluation of Measures. *Human Geography*, 67, 53-67.
- MCDONALD, M. 2012. *Market Segmentation*, Wiley Online Library.
- MULLEY, C., NELSON, J., TEAL, R., WRIGHT, S. & DANIELS, R. 2012. Barriers to implementing flexible transport services: An international comparison of the experiences in Australia, Europe and USA. *Research in Transportation Business & Management*, 3, 3-11.
- O'SULLIVAN, S. & MORRALL, J. 1996. Walking distances to and from light-rail transit stations. *Transportation Research Record: Journal of the Transportation Research Board*, 1538, 19-26.
- REID, R. D. & BOJANIC, D. C. 2009. *Hospitality marketing management*, John Wiley and Sons.
- SHAO, C., XIA, J. C., LIN, T. G., GOULIAS, K. G. & CHEN, C. 2015. Logistic Regression Models for the Nearest Train Station Choice: A Comparison of Captive and Non-captive stations. *Case Studies on Transport Policy*.
- SHEKHAR, S., LU, C.-T. & ZHANG, P. 2003. A unified approach to detecting spatial outliers. *GeoInformatica*, 7, 139-166.
- SMITH, W. R. 1956. Product differentiation and market segmentation as alternative marketing strategies. *Journal of marketing*, 21, 3-8.
- TAYLOR, P. J. 1973. A New Shape Measure for Evaluating Electoral District Patterns. *The American Political Science Review*, 67, 947-950.
- THÉRIAULT, M., CLARAMUNT, C. & VILLENEUVE, P. Y. A spatio-temporal taxonomy for the representation of spatial set behaviours. *Spatio-temporal database management, 1999. Springer*, 1-18.
- WEDEL, M. & KAMAKURA, W. A. 2012. *Market segmentation: Conceptual and methodological foundations*, Springer Science & Business Media.
- WEINSTEIN, A. 2013. *Handbook of Market Segmentation : Strategic Targeting for Business and Technology Firms, Third Edition*, Hoboken, Hoboken : Taylor and Francis.

- WITTEN, I. H. & FRANK, E. 2005. *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann.
- ZHAO, F., CHOW, L. F., LI, M. T., UBAKA, I. & GAN, A. 2003. Forecasting Transit Walk Accessibility: Regression Model Alternative to Buffer Method. *Transportation Research Record*, 1835, 34-41.
- ZHAO, H., YAN, X. & GAO, Z. 2013. Transportation serviceability analysis for metropolitan commuting corridors based on modal choice modeling. *Transportation Research Part A: Policy and Practice*, 49, 270-284.