

## EXPLORING THE POTENTIAL OF CROWD SOURCED DATA TO MAP COMMUTER POINTS OF INTEREST: A CASE STUDY OF JOHANNESBURG

T. Moyo<sup>1\*</sup>, W. Musakwa<sup>2</sup>

<sup>1</sup> Department of Operations and Quality Management, University of Johannesburg, Corner Siemert & Beit Streets, Doornfontein 0184 Johannesburg, South Africa- thembanijoel@gmail.com

<sup>2</sup> Department of Town and Regional Planning, University of Johannesburg, Corner Siemert & Beit Streets, Doornfontein 0184 Johannesburg, South Africa - wmusakwa@uj.ac.za

**KEY WORDS:** Commuter, Johannesburg, Mobility, Crowd sourced data, Demand and Supply

### ABSTRACT:

Modern African cities are faced with various mobility and transportation challenges. In developing smart sustainable cities, city planners need to create a balance between supply and demand for public transportation. Development of multi-mobility mode models has contemporarily received a special interest in smart cities development. Globally, the use of bike sharing services to complete the first kilometre or last kilometre of the trip has been highly received, with commuters using either rail or road mobility modes for the middle section of their trip. Within the developing world context, the use of multi-mobility modes in daily commuting is still new, and little research has been done to guide this. Notwithstanding the influence of uncertainties and fragmentation over demand and supply in public transportation provision. In the South Africa context, various modes of public transportation have been developed which seek to be smart, sustainable and efficient such as the fast train (Gautrain), Bus rapid transport (Rea Vaya and Gaubus) and Bikes sharing platforms (Upcycles), however most of these modes are currently not spatially connected. Hence the researcher sought to develop a stepping stone in planning for future mobility demand. Using an explorative methodology, the authors collected quantitative and spatial data in the form of land-use data and crowd sourced data (from twitter) to map commuter points of interest in and around the city of Johannesburg. The results reveal hot and cold spots in the city. The hot spots reveal areas where commuters frequently travel to, and when overlaid with transportation data, we are able to identify potential locations to develop new transportation hubs as these will overtime become key points of interest.

### 1. INTRODUCTION

Modern African cities are faced with various mobility and transportation challenges. Mobility in urban spaces is continuously transforming and adapting to daily challenges, as can be seen with the global shift towards smart city planning. This shift towards encouraging citizens to utilise multi-mobility modes in their daily commute has led to a reduction in road congestion (Nunes et al., 2014). However despite the benefits, there is still a huge dependence on private vehicles. Scholars have highlight that the aim is not to prohibit the use of private vehicles but to create an efficient public transportation system which would lead to citizens not using private vehicle (Brakewood et al., 2014).

Currently the traditional public transit in Johannesburg is defined by most commuters as being unreliable and not spatially integrated, mainly because of hindrances such as overcrowding and lack of connectivity between the various modes. Literature has identified that improving connectivity in public transit will lead to a sustainable public transportation systems (Moyo & Musakwa, 2016). Thus investment into multi-mobility systems seems to be the best solution to the current public transit problems as scholars have outlined that 'mobility as a service' should prioritise accessibility, whilst also ensuring flexibility between travel modes for various users (Jackiva et al., 2017).

### 2. RELATED WORK

#### 2.1 Crowdsourcing, Internet of Things and Volunteered Geographical Data

Leveraging data from citizens to guide were and how to develop has been gaining attention over the past recent years as researchers and practitioners are now able to mine data from social media platforms (Hasan & Ukkusuri, 2014; Di Pietro et al., 2015; Moyo & Musakwa, 2016; Fontes et al, 2017). Given how most smartphones are equipped with sensors, social media platforms have developed applications which promise to enhance the quality of life of the user by collecting locational data, images and sound in order on the personalise the users experience on the social media platform. For researchers and practitioners this has presented an untapped data source to guide city development.

In a similar fashion, a new term has been coined being the internet of things (IoT) which describes how devices are now connected to every aspect of our lives. IoT has assisted transportation planners to track mobility patterns in the city in real-time, predict future transportation congestion trends, monitor traffic flow within the city (Moyo & Musakwa, 2016). Therefore, as the internet is now entangled in to the citizen's lives as citizens carry their smart phones everywhere, this presents new possibilities for city planning, as mobility information can be collected at real-time and future trends can be predicted.

---

\* Corresponding author

Over the recent years the nature of geographic information has rapidly changed from the traditional cartography legacy, as crowdsourcing and the IoT have change the nature in which location data is collected, stored, processed and visualised (Fontes et al., 2017). A notably example is how users can now use platforms such as Google Maps to search not only for traditional location information but other text searches, (such as food, gym, school and companies), Google Maps will then reveal results with geotagged references. Table 1 summarises

how social data has been used for different research to solve various city scenario problems. In our own study we then explore the potential of assessing commuter movement patterns through the collection of geo-tagged tweets (from twitter social media) concerning the two leading public transportation systems in Johannesburg namely Rea Vaya and the Gautrain (also the Gaubus a bus rapid transit system which acts as an extension of the Gautrain).

Author(s)	Evaluation Approach	Results	Case Study
Fontes et al., 2017	Mobile payments	The survey revealed that commuters valued convince and time saving when making payments for their daily trips.	Oporto City and Beijing City
Brakewood et al., 2014	Mobile payments	The results from the study illustrates noteworthy the potential for adoption of mobile ticketing on the commuter rail network. Mobile ticketing offers commuters a more convenient purchase method than prepayment at ticket windows or vending machines and is less problematic (for both the customer and operator) than on board cash transactions. In light of this study and the widespread adoption of smartphones, mobile ticketing appears to be a compelling alternative to traditional ticketing methods, and its adoption by rail operators and utilization by riders are likely to increase in the near future.	Boston
Di Pietro et al., 2015	Mobile payments	The results illustrated how the objective to use an Integrated Model on Mobile Payment system is influenced by the Usefulness, Ease of use and the Security of that payment system. Furthermore, the Usefulness is concurrently influenced by the Ease of use, the Compatibility with users' values and needs and their Attitude towards mobile services.	Italy
Chowdhury & Ceder, 2016	Route planner	Easy of transferability between various modes is essential for an integrated transport system	Literature Review
Jackiva et al., 2017	Route planner	Public Transport provision should be integrated from the inception stage to ensure sustainability at various spatial levels for local and long distance trips	Riga City
Moyo & Musakwa, 2016	Web 2.0	If web 2.0 with spatio-location records is mined, commuter experiences of the city can be visualised. However for the data to be useful, there is need for mechanisms to validate the big data.	Johannesburg City
Hasan & Ukkusuri, 2014	Web 2.0	Results from their study reveal that data which is spatial and temporal has the prospective of mapping out the movement patterns of the people at either a global or local scale. However while data generated automatically generally has limitations, the sheer granularity has the potential to open new research avenues, which were previously unattainable from data acquired from field surveys.	New York City

2.2 Table.1: Summary of literature review on Research on Crowdsourcing, Internet of Things and Volunteered Geographical Data

### 2.3 Connectivity

Contemporary trends in most transportation systems have been shaped by demand and supply constraints. This has led to city planners seeking to visualise connectivity in various mobility networks in order to balance supply and demand. In transportation planning, the concept of connectivity was has been used to evaluate the importance of a mobility mode in terms of its connection to other mobility networks. Scholars

have developed various connectivity measures, to visualise the degree to which mobility modes are connected (Zhu, Zhang & Zhang, 2018). Given the relationship between transportation, connectivity and economic development, different connectivity measures have been developed using various constraints such as trip purpose, trip duration, and many others as a means to ensure objective in the analysis (Mishra, Welch & Jha, 2012; Cheng & Chen, 2015).

However most studies only focus on connectivity of one mobility mode, whilst in reality there exists several main mobility modes in the city, which complement each other. Hence this study shall seek to visualise this connectivity by focusing on Bus rapid transit systems and railway systems.

### 3. METHODOLOGY

#### 3.1 Study Area

The research shall focus on mobility networks within the city of Johannesburg (see Figure 1). Johannesburg is situated within

the Gauteng province, and this city has been coined the economic hub of the province due to the various economic activities in this city. This has led to many citizens migrating toward the city in search of greener pastures. Although the city is land-locked, it is the faster growing meg-city in Africa. This has led to the existing mobility no longer being able to cater for the increased population, hence a call to explore how to improve the existing connectivity between existing transportation modes to ensure free-flow of commuters during their commuting trips.

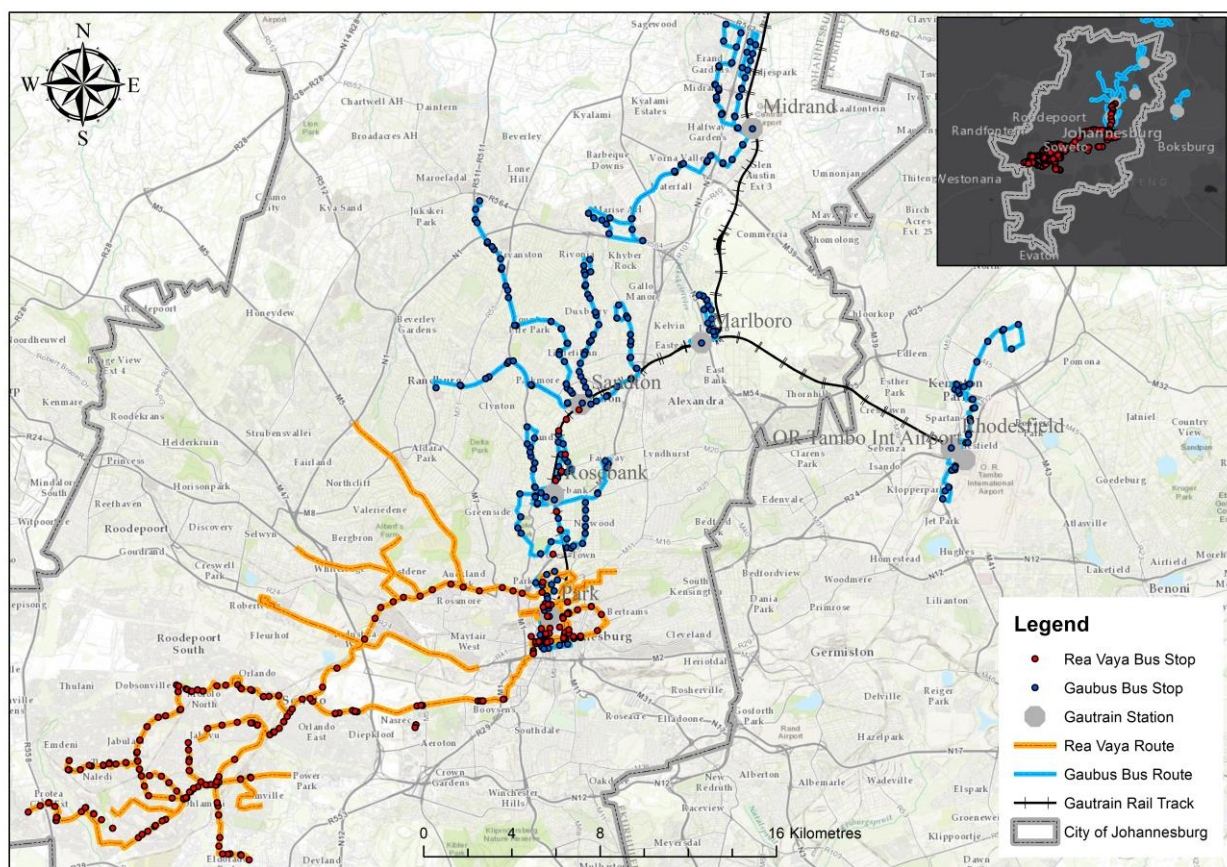


Figure 1: Study Area

#### 3.2 Data Collection and Analysis

Using an explorative methodology, the authors collected quantitative and spatial data in the form of land-use, big data and crowd sourced data (twitter) to map commuter points of interest in the city of Johannesburg. Figure 2 summaries how the data was collected and how the preliminary analysis was carried out to ensure only tweets with geo-location data about the two mobility modes (Gautrain and Rea Vaya) were used in the analysis. After preliminary analysis 11000 tweets (with geolocation properties) were collected over a 6 month period from June to December 2018 using the key words: 'Gautrain'; 'Gaubus'; and 'Rea Vaya'.

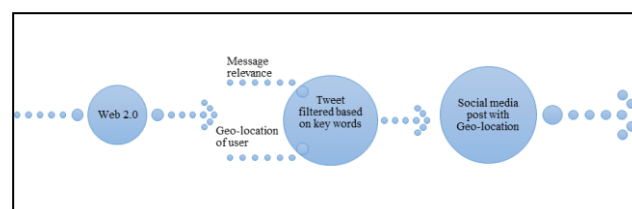


Figure 2. Preliminary analysis

The neighbourhood analysis tool in Arc Gis point statistics estimates attributes such as what is adjacent or nearby of a location using data from surrounding points. In this research, we utilised this Arc Gis tool to visualise the geographic concentration of commuters of the two mobility modes. This allowed us to assess the relationship between their locations, available transportation infrastructure and land use in their proximity. In research proximity is referred to as the capacity to



classify the spatial distance between of parcel features. The social media data from twitter was overlaid to assess the spatial pattern variation that existed in the City of Johannesburg.

#### 4. RESULTS AND DISCUSSION

The city of Johannesburg is currently faced with a challenge of continuing to be the economic hub of the Gauteng province whilst still meeting the mobility needs of commuters. Creation of new transportation infrastructure may prove to be an expensive endeavour, hence the need to ensure synergies are shared by the existing mobility systems. This could prove beneficial for both commuters and transportation providers as for commuters their mobility needs will be met, and for the service providers having more commuters

For the study the authors initial focused on visualising the commuter concentration density based the locations commuters posted from (See figure 4). The highest commuter concentration ranging from 8500 to 11 000 users were located near existing mobility hubs such as Park, Rosebank, and Sandton. When overlaid with land-use data these locations are within close proximity to areas of economic opportunities and business hence commuters traverse to the locations frequently. Trips made to these locations are hence commuting trips, meaning they occur on a daily basis. With this knowledge, there is a need for city planners to invest more into improving mobility modes

within this vicinity as this will led to economic benefits for the city and country at large.

Furthermore locations on the north-western part of the city such as Fourways and Randburg also have a high commuter concentration. The locations have office parks and residence areas. Trips made to these locations are either for commuting or leisure purposes. A further content analysis of the social media posts made in this location, reveals that commuters are concerned with travel time. City planners could use this guide the development of dedicated public transportation routes such as for the BRT system. In investment into a reliable public transportation system will lead to sustainable developmental practises.

With this in mind the authors conducted a sematic analysis to determine commuter opinions regarding the modes of public transportation namely the Gautrain and Rea Vaya. (See figure 5). From the analysis of the most frequent used words in the posts, commuters prioritize the following ‘time’, ‘movement’ and ‘efficiency’. Hence to have a surge in commuter volumes using public transportation, there is a need to ensure the existing mobility modes meet these requirements. Also the occurrence of delays within a public transportation system would lead to adverse effects.

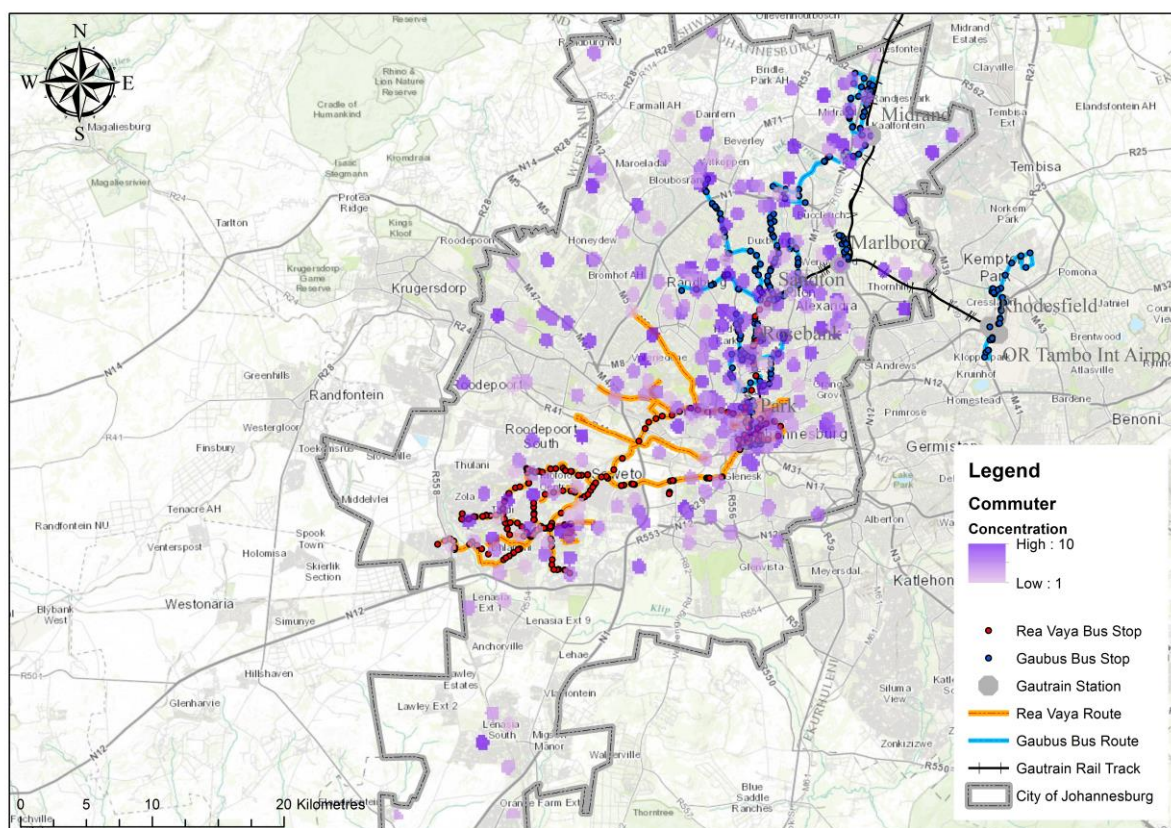


Figure 3. Commuter concentration

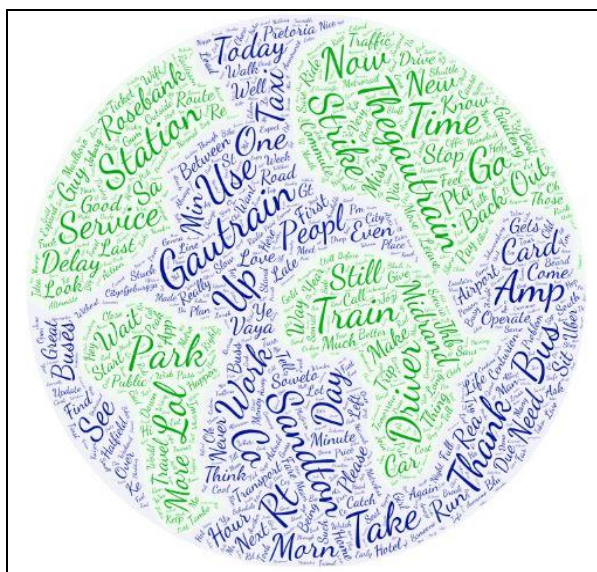


Figure 4. Frequently used words in Tweets during 2018

For the six month period in 2018, the highest numbers of trips both commuting and recreational occur in the during the morning peak for both modes of public transportation. This could be due to the commuters' need for reducing the total travel time when going to work, school or making trips for administrative duties. From the content analysis, most of the posts made during this time are complaints are delays in travel time.

From 12pm to 4pm, there is an increase in commuting trips, this trips could be for leisure and commuting purposes. In the city of Johannesburg there is usually an increase in traffic congestion during these times. This correlates to the results obtained from twitter, as most of the posts, related to traffic accidents on various major roads, and delays on the freeway. After 4pm there is a decrease in the number of posts. This could be due to having less commuters travelling during this time of day, as most commuters leave work at 4pm.

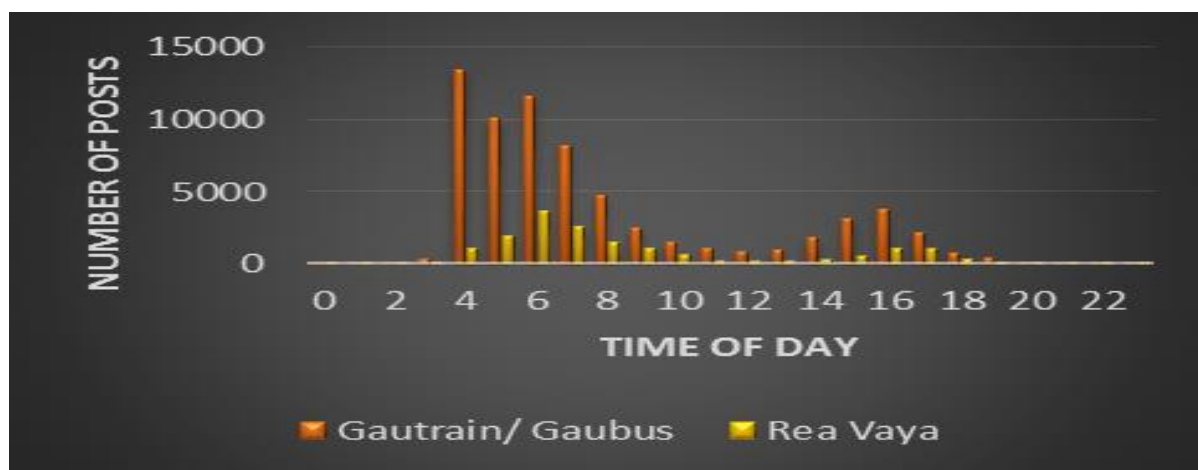


Figure 5. Concentration of trips per hour

## 5. CONCLUSION

The key findings from the research is that social media data can be used to visualise points of interest of commuters. Also the socio-economic implications to transportation planning are evident, in the dynamics of tweets posted. That is service providers can use social media data to predict mobility trends, this in turn will lead to less negative posts being made by users. Once commuters feel their concerns are being addressed by the service provider, they will reciprocate by use that mode of transportation more often and encourage the immediate social network to also use this mode of transportation.

With regards to future work, the authors propose incorporating data about other public transportation systems such as mini-bus taxis as these currently have a larger route network, and their integrated into mobility planning for the future presents an untapped connectivity network.

## ACKNOWLEDGEMENT

The research in this paper was funded by the National Research Fund (NRF), South Africa, Grant number 110778.

## REFERENCES

- Brakewood, C., Roja, F., Robin, J., Sion, J., Jordan, S. 2014. Forecasting mobile ticketing adoption on commuter rail. *Forecasting Mobile ticketing adoption on commuter rail. J. of Public Transportation*, 17, pp.1–19.
- Cheng, Y.H. and Chen, S.Y., 2015. Perceived accessibility, mobility, and connectivity of public transportation systems. *Transportation Research Part A: Policy and Practice*, 77, pp.386–403.
- Chowdhury, S. and Ceder, A.A., 2016. Users' willingness to ride an integrated public-transport service: A literature review. *Transport Policy*, 48, pp.183–195.

Di Pietro, L., Mugion, R.G., Mattia, G., Renzi, M.F. and Toni, M., 2015. The integrated model on mobile payment acceptance (IMMPA): an empirical application to public transport. *Transportation Research Part C: Emerging Technologies*, 56, pp.463-479.

Fontes, T., Costa, V., Ferreira, M.C., Shengxiao, L., Zhao, P. and Dias, T.G., 2017. Mobile payments adoption in public transport. *Transportation research procedia*, 24, pp.410-417.

Hasan, S. and Ukkusuri, S.V., 2014. Urban activity pattern classification using topic models from online geo-location data. *Transportation Research Part C: Emerging Technologies*, 44, pp.363-381.

Jackiva, I. Y., Budiloviča, E.B., Gromule, V. 2017. Accessibility to Riga Public Transport Services for Transit Passengers. 10th International Scientific Conference Transbaltica 2017: Transportation science and Technology. *Procedia Engineering* 187 (2017) 82 – 88

Mishra, S., Welch, T.F. and Jha, M.K., 2012. Performance indicators for public transit connectivity in multi-modal transportation networks. *Transportation Research Part A: Policy and Practice*, 46(7), pp.1066-1085.

Moyo, T & Musakwa, W (2016). Using crowdsourced data (Twitter & Facebook) to delineate the origin and destination of commuters of the Gautrain public transit system in South Africa. *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences*, Volume III-2, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic. doi:10.5194/isprsannals-iii-2-143-2016. pp 150

Nunes, A. A., Galvão, T., Falcão e Cunha, F. 2014. Urban public transport service co-creation: leveraging passenger's knowledge to enhance travel experience. EWGT2013 – 16th Meeting of the EURO Working Group on Transportation. *Procedia - Social and Behavioral Sciences* 111 (2014) 577 – 585

Yang, H.C. and Lee, C.H., 2013. Toward crowdsourcing data mining. In *The 3rd International Workshop on Intelligent Data Analysis and Management* (pp. 107-110). Springer, Dordrecht.