

THE USE OF GENERAL TRANSIT FEED SPECIFICATION (GTFS) APPLICATION TO IDENTIFY DEVIATIONS IN THE OPERATION OF PUBLIC TRANSPORT AT MORNING RUSH HOUR ON THE EXAMPLE OF SZCZECIN

Sławomir Goliszek¹, Marcin Połom²

¹ Institute of Geography and Spatial Organization

Polish Academy of Sciences

Twarda 51/55, 00-818 Warsaw: Poland

sgoliszek@twarda.pan.pl

² University of Gdansk

Bażyńskiego 4, 80-309 Gdansk: Poland

marcin.polom@ug.edu.pl

Abstract. The article examined the possibility of using General Transit Feed Specification (GTFS) to identify deviations of public transport in morning hour traffic. Deviations in the functioning of public transport spanned in a radius of 30 minutes public transport travel time from the center of the city. The adopted travel time to the center is taken from a comprehensive traffic study performed in the city where indicated value of 27 minutes (comprehensive traffic study ...) is used as the average travel time of public transport. Diagnosis of deviations in the operation of public transport was taken on a weekday between 6:30 and 8:30 am at 5 min intervals. The results of calculations of time availability were made every 5 minutes and applied to each other and trimmed among themselves. Hence the contour line was established within isochrone of 30 minutes, with the identification of areas with 100% certainty of public transport access to the center of the city and areas of lesser public transport accessibility. The resultant isolines were measured against the population living in census districts, on the basis of which the areas of certain as well as uncertain 30 minute travel time to the center of the city were determined.

Keywords: General Transit Feed Specification (GTFS), public transport, isochrones.

Introduction

Studies on the availability of public transport systems in urban settlements have been conducted for many decades and based mainly on indicators and characteristics of the public transport functioning (Sobczyk 1985). Analyses of public transport in cities in most cases include valuable investments in this area. In recent years, public transport in Poland has been supported by structural funds obtained from the EU. Thanks to funding from the European Union's 2004-2006, 2007-2013 and currently

ongoing 2014-2020 financial perspectives made the investments possible. Without that support it would be difficult to implement such actions. In the last two perspectives an increasing number of projects were conducted to improve transport accessibility in urban settlements. Regional centers of Eastern Poland received additional measures to improve the functioning of public transport. Some of these projects implied establishing the new means of transport (Goliszek 2014; Goliszek, Rogalski 2014). In contrast, the other investments included comprehensive support action, involving the purchase of new rolling stock, deployment of ITS and the development of existing infrastructure (Połom & Tarnawski 2011; Goliszek 2014b, 2014c).

GTFS format was created in the United States of America (USA) in the early 1990s. Initially, the patronage of the introduction of the format of the data has been coordinated by Google. Currently the tools to use GTFS data can be found in ESRI's ArcGIS. This format of public transport coding data has been currently used by most boards of public transport companies in the US and in Western and Central Europe. Hence research on public transport in these areas is measured with high precision (Widener et al. 2015; Farber et al. 2014). GTFS data format was also used for comparisons of public transport systems (Poelman et al., 2015), and modeling travel characteristics based on the selected mean of transportation or lack of it. A large part of the studies using GTFS is utilized to compare public and individual transport (Salonen et al. 2013; Wang et al. 2015). Comparison of individual and public transport can lead to conclusions on their functioning and how to promote intelligent transport systems in urban areas (Malecki et al. 2014; Tao et al. 2014). The creation of General Transit Feed Specification (GTFS) data format and their implementation by the authorities responsible for public transport in the city created new possibilities for the analysis of transport. GTFS data format is more often used to model the transport accessibility (rail, tram, trolley, bus, aviation, regional and automotive). Therefore it is possible to study multimodal transport accessibility by implementing the modeling of all means of transport (Hadas 2013; Hadas et al. 2012).

The main objective of the study is to investigate the deviations in the operation of public transport using the tools and GTFS data format. The study of public transport accessibility is a long known idea. Until now, the passengers were surveyed for their satisfaction with the functioning of public transport (Mouwen 2015) and the impact of tourism on the functioning of urban public transport (Lapko 2014). Some studies focused on accessibility to public transport stops or their comparisons between different urban centers (Fiori et al. 2016; Falavigna et al. 2016). At the same time it should be noted that the study of the frequency of public transport courses and the exact demarcation of the areas of transport accessibility of public transport are the new fields of study. Currently, computer equipment allows researchers to make enormous matrix calculations and research the availability of the communication network, analyses that were previously difficult.

Data

The primary source of data used in the analysis is the General Transit Feed Specification for the city of Szczecin. These data contain information about the location of stops, frequency of public transport service and directions of public transport transit lines. The Roads and Transport Management (ZDiTM) in Szczecin provides data in GTFS format. Data are ready for use in the ArgGIS 10.2 tool Network Analyst. An important source of data that were used in the analysis to identify deviations in the operation of public transport is certain information from a number of populations divided into statistic perimeters (2011). These data are available to the public and have been downloaded

from the geostatistical portal. Additional data sources are obtained by Open Street Map saved as shp format, which were used to generate the connections between stops as paths of pedestrian crossing to / from places where public transport does not reach (the average walking speed of 4.5 km/h; Farber et al. 2014; Ratajski 1989).

Methods

To identify deviations in the functioning of public transport using the GTFS application, the qualitative methods were applied in an analysis. This method is based on a detailed positioning of isochrones of access to the center in 5 minute intervals. The test was performed within two hours during morning rush hours. The method used in the analysis is the isoline of time (isochrone). The results for all the time intervals at the morning peak were trimmed to each other and a core of 100% certainty access within 30 min to the center of the city at morning rush hours was established (Warakomska, 1970). However, all isolines of access to the center beyond 30 minutes characterized points of uncertainty which allowed identification of deviations in the functioning of public transport in peak (Sierpinski 2010). During the identification process, calculation of standard deviation proved to be helpful in 10 minute intervals (0-10; 10-20; 20-30; 30-40; 40-50; 50-60; 60-70; 70-80; 80-90 and for each isoline in 5 minute intervals).

Analysis of the population in isochrones

Analysis of the population marks isolines in a maximum range of 90 min travel time using public transport to the center of Szczecin was performed. The total sum of the number of population living in the area of 90 minute travel time using public transport is approximately 470 thousand people (+/- 1%), which means that the public transport system organized by ZDiTM in Szczecin allows almost all the inhabitants of the entire city as well as part of the area located outside metropolis to travel.

The results of the analysis of 25 measurements taken, indicate obvious deviations in the operation of public transport for the passengers traveling by public transport. Most noticeable deviations reside in those sections of the population that are located in the isochrone with a percentage of 20 to 90% of the population (470 thousand. = 100%). The participation of inhabitants of Szczecin in access isolines indicates high fluctuations in travel time up to 20 minutes towards the center of the city. In the isochrones up to 20 minutes, even small fluctuations in the functioning of public transport cause a major decrease in the availability of transport for several thousand people (Fig. 1). Such high variation between the different sections explains wide variety of the public transport network.

Summary of the results in tabular form enabled indication of the periodicity or "temporality" in the functioning of public transport. Results of periodicity of the population showed in isochrones indicate evenly distributed maximum and minimum values. Data in sections are arranged at an angle of 45 degrees, which means that the maximum value is repeated in the next interval, but for larger values of the travel isochrone. On the other hand an analysis of the surface at intervals of isochrones indicates the concentration of results in certain sections of travel time.

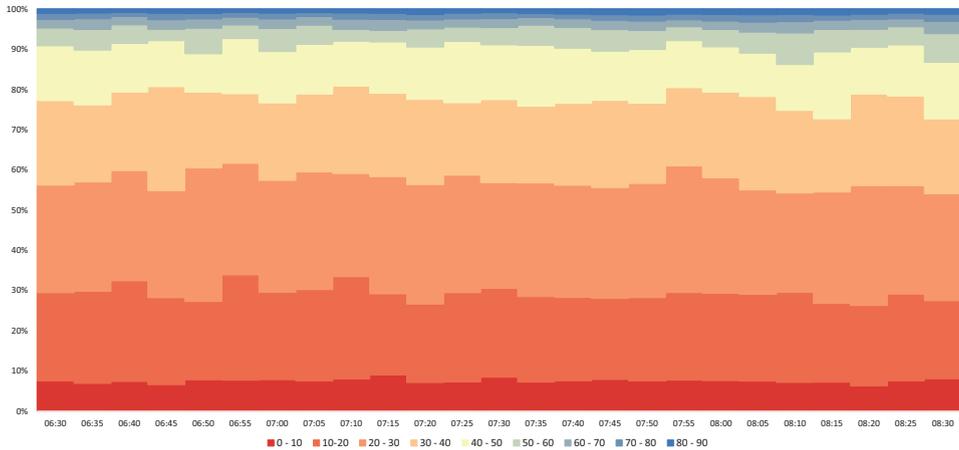


Figure 1. The percentage structure of population number in the analysis of 5 minute intervals

Comparison of the simulation results at the peak of transport hours provides information about the extent of deviation from the mean (standard deviation). At intervals, the smallest standard deviation is in the range of 0-10 minutes, where the result of this effect is caused by a low level of use of public transport, and most of the travel is on foot at the average speed of 4.5 km/h (Farber et al. 2014; Hadas 2013). A high value of the standard deviation is noticeable at isochrone of 50 to 60 minutes. Within this time covering the distance, public transport network is significantly branching along the main roads. Places located outside of Szczecin are covered within a time exceeding 50 minutes. The fact that the city is located on both sides of the river (Odra) which acts as a natural transport barrier, is also of importance. Getting to the other side of the river to the east side of the city is possible in two places, making it difficult to integrate the city and the functioning of public transport. The low value of deviation in the range of over 80 minutes is caused by low population density. The situation is reversed in intervals of 10 to 40 minutes of travel time, where high population density and high variability of the availability of public transport gives high values.

Table 1. The percent of city Surface and population number in intervals of isochrones during morning rush hours

min/time	06:30	06:35	06:40	06:45	06:50	06:55	07:00	07:05	07:10	07:15	07:20	07:25	07:30	07:35	07:40	07:45	07:50	07:55	08:00	08:05	08:10	08:15	08:20	08:25	08:30	stand. deviation
0 – 10	0,12	0,11	0,12	0,09	0,14	0,12	0,11	0,13	0,13	0,15	0,12	0,12	0,14	0,12	0,13	0,13	0,13	0,14	0,14	0,14	0,12	0,11	0,10	0,13	0,14	0,013
10-20	1,37	1,27	1,39	1,21	1,26	1,77	1,30	1,39	1,47	1,32	1,24	1,26	1,32	1,20	1,20	1,21	1,27	1,38	1,31	1,24	1,23	1,27	1,24	1,27	1,23	0,120
20 – 30	3,96	4,08	5,34	4,45	4,68	4,86	4,10	4,50	4,92	4,78	3,82	4,65	3,97	3,40	3,76	4,12	4,15	4,33	4,18	4,03	4,02	4,02	4,22	3,93	3,84	0,437
30 – 40	6,78	6,49	7,96	7,91	7,43	7,85	7,06	7,44	8,15	7,69	7,15	6,49	6,97	6,66	6,65	6,65	6,48	7,61	7,72	7,84	6,68	5,70	7,33	7,53	6,44	0,626
40 – 50	11,4	10,9	12,6	12,6	10,6	14,7	11,6	13,0	13,7	13,3	11,5	10,9	11,6	12,2	11,7	11,3	10,7	11,8	11,1	10,4	9,1	10,1	10,4	12,1	10,5	1,237
50 – 60	16,2	16,0	19,2	17,9	17,9	18,1	17,9	19,6	17,3	16,3	17,3	15,7	15,2	16,5	16,4	16,3	15,5	16,2	15,8	15,3	15,1	13,5	17,4	18,7	14,3	1,494
60 – 70	20,0	21,6	20,9	21,2	18,7	18,7	20,6	20,5	20,0	19,7	19,8	18,5	18,3	18,5	18,4	18,3	17,6	17,9	17,9	17,9	20,1	19,2	20,2	19,3	19,2	1,136
70 – 80	20,0	21,2	20,6	20,6	19,5	18,7	20,8	20,0	19,6	20,6	19,6	17,9	18,1	18,7	18,7	18,8	19,4	19,1	19,1	19,2	19,5	19,7	19,7	19,4	19,4	0,812
80 – 90	20,3	19,7	17,9	17,7	18,7	17,9	19,2	18,1	18,1	19,1	20,0	20,7	19,6	19,3	19,9	20,6	21,7	21,3	20,5	19,7	20,4	21,9	19,9	19,5	21,3	1,201
stand. deviation	8,31	8,634	8,321	8,345	7,944	7,742	8,501	8,359	7,936	8,115	8,299	7,864	7,674	7,949	7,971	8,005	8,122	8,003	7,866	7,735	8,24	8,341	8,322	8,271	8,228	0,523
stand. deviation 30 min	1,96	2,04	2,72	2,27	2,36	2,41	2,05	2,25	2,47	2,41	1,9	2,35	1,96	1,67	1,87	2,06	2,07	2,15	2,08	2,01	2,01	2,01	2,13	1,95	1,9	1,76
							Low % of the value					Average % of the value					High % of the value						
min/time	06:30	06:35	06:40	06:45	06:50	06:55	07:00	07:05	07:10	07:15	07:20	07:25	07:30	07:35	07:40	07:45	07:50	07:55	08:00	08:05	08:10	08:15	08:20	08:25	08:30	stand. deviation
0 – 10	7,44	6,87	7,35	6,57	7,78	7,65	7,79	7,44	7,95	8,96	7,08	7,19	8,40	7,19	7,52	7,85	7,46	7,69	7,55	7,41	7,10	7,14	6,26	7,50	8,04	0,550
10-20	22,0	22,9	25,1	21,6	19,5	26,2	21,8	22,8	25,5	20,2	19,5	22,3	22,1	21,3	20,7	20,2	20,7	21,8	21,7	21,6	22,4	19,6	20,0	21,6	19,4	1,809
20 – 30	26,8	27,2	27,4	26,7	33,2	27,8	27,8	29,3	25,7	29,2	29,7	29,2	26,3	28,3	27,9	27,6	28,4	31,5	28,8	26,0	24,8	27,8	29,8	27,0	26,6	1,844
30 – 40	21,0	19,2	19,6	25,9	18,9	17,3	19,3	19,4	21,8	20,7	21,2	18,0	20,6	19,1	20,4	21,7	20,0	19,4	21,3	23,2	20,6	18,2	22,8	22,3	18,6	1,889
40 – 50	13,6	13,6	12,1	11,4	9,5	13,7	12,8	12,4	11,2	12,8	13,0	15,3	13,6	15,2	13,7	12,2	13,4	11,7	11,3	10,8	11,4	16,6	11,6	12,7	14,1	1,571
50 – 60	4,42	5,13	4,63	2,85	6,35	3,40	5,76	4,75	2,95	2,90	4,52	3,55	4,31	5,02	5,19	5,36	4,73	3,45	4,34	5,23	7,85	5,60	4,49	4,53	7,15	1,235
60 – 70	2,13	2,68	2,08	2,37	2,30	1,91	2,35	2,15	2,52	2,70	2,22	2,01	2,23	1,93	2,13	2,33	2,34	1,71	2,09	2,47	2,85	2,30	2,46	1,95	3,06	0,311
70 – 80	1,50	1,40	1,02	1,60	1,37	1,20	1,44	1,00	1,55	1,53	1,38	1,42	1,45	0,97	1,19	1,40	1,49	1,43	1,65	1,87	1,77	1,29	1,45	1,25	1,00	0,228
80 – 90	1,05	0,97	0,80	1,01	1,06	0,80	0,98	0,82	0,95	1,03	1,30	1,01	0,87	1,07	1,21	1,34	1,46	1,23	1,30	1,37	1,29	1,45	1,25	1,00	1,26	0,200
stand. deviation	10,01	9,952	10,48	10,77	10,78	10,68	9,846	10,48	10,47	10,25	10,33	10,43	9,878	10,12	9,962	9,86	9,96	10,91	10,34	9,891	9,228	9,644	10,62	10,16	9,009	0,742
stand. deviation 30 min	10,1	10,7	11	10,5	12,7	11,2	10,3	11,2	10,2	10,1	11,3	11,3	9,37	10,7	10,4	9,97	10,6	12	10,8	9,72	9,58	10,4	11,8	10,1	9,37	9,26
							Low % of the value					Average % of the value					High % of the value						

Results of commuting analysis [30 min]

For a better perception of development of the situation in terms of commuting to the city center the highest isolines were selected, characterized by the greatest population number, which is presented on the map. The largest isochrones directions to the city center in the range within 30 minutes do not reflect the range of most surface baseline value. The network of the public transport is easily identifiable on the map. The isochrones distribution is very similar to the one on the maps, however choosing the best isochrone is determined by the transport distribution in Szczecin, which is correlated with public transport points (Fig. 2). By far, the greatest impact on lengthening or shortening of isochrones is assigned to the currently operating system of tram lines complemented by bus network.

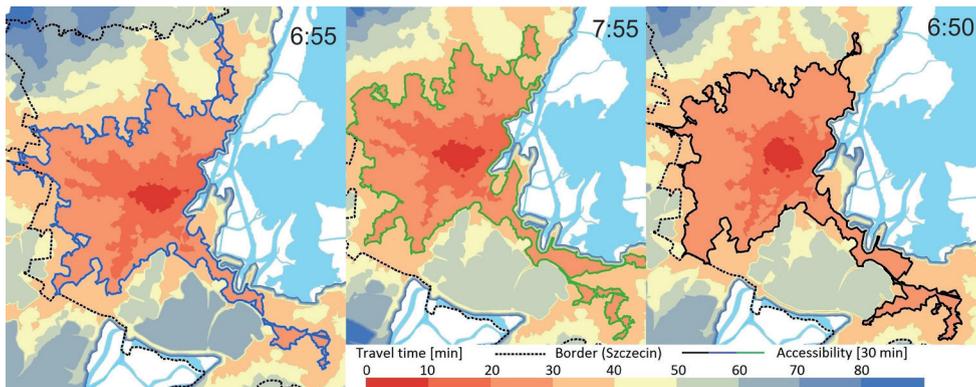


Figure 2. An example of the most efficient time availability in isochrone up to 30 minutes for the number of population (1-left, 2-middle, 3-right)

The results of a detailed analysis of „certain” and „uncertain” travel

The results of a detailed analysis of 25 images transport availability up to 30 min to the center differ from each other. On the basis of detailed analysis, it is possible to mark the area where the inhabitants of Szczecin have a “certain” access to the city center in the morning rush hours. Development of a “certainty” map enabled designation of a number of “uncertain” access points for travel to the center using public transport. The whole area within which it was possible to reach the center in 30 minutes included the total number of 329.1 thousand people as inhabitants. 214.4 thousand inhabitants have been in range of “certain” access to public transport. The population living in a separate area of guaranteed travel access on the map (red color) can be confident that public transport travel times in their case is less than 30 min (Fig. 3). The “uncertain” travel access area, where it is still possible to reach the center in 30 minutes however under condition of reaching the stop at the right moment, is inhabited by 114.7 thousand people. The ratio between the area of “certain” and “uncertain” access is 65:35. Among those with guaranteed time travel in 30 minutes to the center, 100% population consists of residents of Szczecin (52.7% of the population). However, in area where the travel time is not guaranteed under 30 minutes to the center, more than 95%

people are the residents of Szczecin (28.2% of the population), and 5% are residents of the areas outside the city limits (Fig. 3).

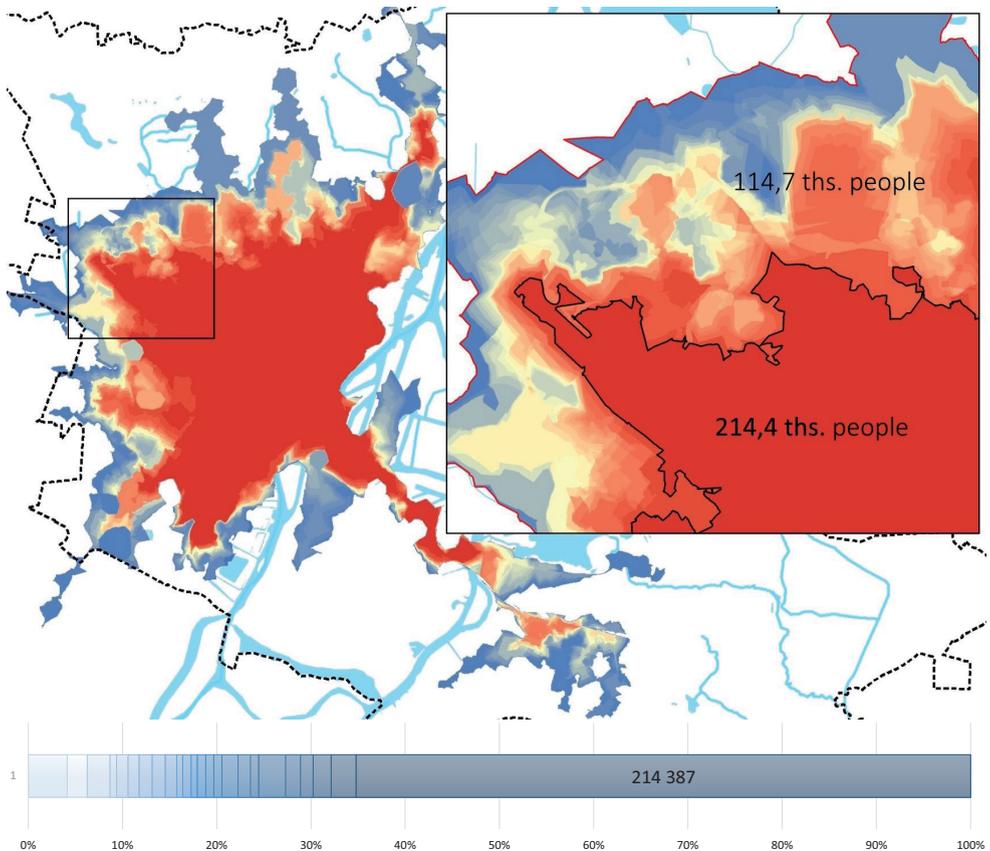


Figure 3. Distribution of 30 minutes isochrone to the city centre and number of population residing in this area

When analyzing in detail the number of people living in the area where access is not guaranteed, the largest part of the population has 8% chance to reach the center in rush hours, which translates into low likelihood of travel for more than 13.6 thousand people. Another area with uncertain travel access is inhabited by 9.3 thousand people, and the population living there has a high 80% probability of access. In places located in close proximity to the area of “certainty” population of 9 thousand people have a very high chance of access at the level of almost 96 percent. These results indicate a high fluctuation of distribution of the population in the “uncertain” areas of the city. Better reception of studied matter presents itself in the population density in the area of uncertain access to the city center. Population density in the studied area of uncertain travel access to public transport is distributed linearly of approximately 4 thousand/sq km in high “certainty”, up to approximately 2 thousand/sq km in the area of high “uncertainty” of travel access. In relative values, the areas located closer to the center are more densely populated, while the areas of lower density are more remote. However, in the absolute values the number of population of the sections varies, so that

the highest values occur at the beginning and end of the section of “uncertainty” in travel access. The only exception, which disrupts the linear graph of population density, is the place of 80% access guarantee, which is inhabited by 6.1 thousand/ sq m and differs significantly from the border areas of population density in the “uncertain” travel areas (Fig. 4).

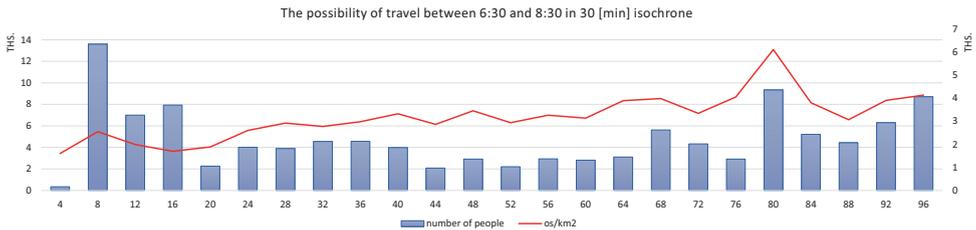


Figure 4. Population density in the area of „uncertain” travel with reference to the number of population in certain intervals

Summary

The use of applications and data in the GTFS format to identify deviations in the functioning of public transport gives sound results. Data provide detailed qualitative and quantitative information, whereas cutting them to the population number in the census circuits provides accurate population data. The results clearly indicate large variations in the availability of public transport. The area of “certain” public transport access is inhabited by more than 50% of the inhabitants of Szczecin. In this area, variability in the functioning of public transport depends on overlapping schedules (both bus and tram). In the area of “uncertain” travel access to the center of the city, public transport variability can be associated with functioning of the bus network. The area of “uncertain” access is inhabited by approximately 25% of Szczecin residents.

GTFS data format enables identification of public transport functioning, provides directions of overdemand or oversupply in the functioning of public transport and detailed identification of transport during rush hours and beyond. Data encoding public transport and creation of a common format for all transport systems facilitates the implementation of in-depth analysis and identification of errors in timetables and adapting them to the needs of travelers. GTFS analytic tools can be a great asset to support the management boards of regional and municipal public transport.

References

- Aarhaug J., Elvebakk B., 2015. *The impact of Universally accessible public transport—a before and after study*. Transport Policy, vol. 44, pp. 143–150.
- Falavigna C., Hernandez D., 2016. *Assessing inequalities on public transport affordability in two latin American cities: Montevideo (Uruguay) and Córdoba (Argentina)*. Transport Policy, vol. 45, pp. 145–155.
- Farber S., Morang M. Z., Widener M. J., 2014. *Temporal variability in transit-based accessibility to supermarkets*. Applied Geography, vol. 53, pp. 149–159.

- Fiori A., Mignone A., Rospo G. 2016. *DeCoClu: Density consensus clustering approach for public transport data*. Information Sciences, vol. 328, pp. 378–388.
- Fransen K., Neutens T., Farber S., De Maeyer P., Deruyter G., Witlox F., 2015. *Identifying public transport gaps using time-dependent accessibility levels*. Journal of Transport Geography, vol. 48, pp. 176–187.
- Goliszek S., 2014a. *Poprawa dostępności miejskim transportem zbiorowym w Olsztynie w świetle inwestycji infrastrukturalnych z perspektywy UE 2014-2020*. Transport Miejski i Regionalny, vol. 5, pp. 30-36.
- Goliszek S., Rogalski M., 2014. *Przestrzenno-czasowe zmiany dostępności komunikacyjnej miejskim transportem zbiorowym w Rzeszowie w świetle inwestycji współfinansowanych ze środków UE na lata 2014-2020*. Transport Miejski i Regionalny, vol. 7, pp. 23-30.
- Goliszek S., 2014b. *Zmiany dostępności miejskim transportem zbiorowym w Lublinie w wyniku inwestycji infrastrukturalnych finansowanych z funduszy UE do roku 2020*. Transport Miejski i Regionalny, vol. 9, pp. 15-21.
- Goliszek S., 2014c. *Dostępność komunikacyjna transportem zbiorowym w Białymstoku – wpływ środków z perspektywy UE na lata 2014-2020*. Transport Miejski i Regionalny, vol. 11, pp. 19-26.
- Hadas Y., 2013. *Assessing public transport systems connectivity based on Google Transit data*. Journal of Transport Geography, vol. 33, pp. 105–116.
- Hadas Y., Ranjitkar P., 2012. *Modeling public-transit connectivity with spatial quality-of-transfer measurements*. Journal of Transport Geography, vol. 22, pp. 137–147.
- Gmina Miasto Szczecin, 2010. *Kompleksowe Badania Ruchu w Szczecinie w 2010 roku*. Szczecin.
- Krawczyk G., 2013. *Strategiczne zarządzanie rozwojem transportu zbiorowego w Polsce*, Transport Miejski i Regionalny, vol. 2, pp. 3-7.
- Łapko A., 2014. *Urban Tourism in Szczecin and its Impact on the Functioning of the Urban Transport System*. Procedia – Social and Behavioral Sciences, vol. 151, pp. 207–214.
- Małecki K., Iwan S., Kijewska K., 2014. *Influence of Intelligent Transportation Systems on Reduction of the Environmental Negative Impact of Urban Freight Transport Based on Szczecin Example*. Procedia-Social and Behavioral Sciences, vol. 151, pp. 215-229.
- Mouwen A., 2015. *Drivers of customer satisfaction with public transport services*. Transportation Research Part A: Policy and Practice, vol. 78, pp. 1–20.
- Poelman H., Dijkstra L., 2015. *Measuring access to public transport in European cities*. Regional Working Paper.
- Połom M., Tarnawski R., 2011. *Wsparcie modernizacji i rozwoju komunikacji miejskiej w Lublinie z funduszy strukturalnych*. Transport Miejski i Regionalny, vol. 10, pp. 35-41.
- Ratajski L., 1989, *Metodyka kartografii społeczno-gospodarczej*. Warszawa: Państwowe Przedsiębiorstwo Wydawnictw Kartograficznych.
- Salonen M., Toivonen T., 2013. *Modelling travel time in urban networks: comparable measures for private car and public transport*. Journal of Transport Geography, vol. 31, pp. 143–153.
- Sierpiński G., 2010. *Miary dostępności transportowe miast i regionów*. Zeszyty Naukowe Politechniki Śląskiej, vol. 1825, Katowice.
- Sobczyk W., 1985. *Dostępność komunikacyjna w układach osadniczych miast*. Warszawa: Komitet Badań Regionów Uprzemysłowionych.
- Tao S., Rohde D., Corcoran J., 2014. *Examining the spatial-temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap*. Journal of Transport Geography, vol. 41, pp. 21–36.

- Todd L., 2013. *The New Transportation Planning Paradigm*. ITE Journal, vol. 83, pp. 20-28.
- Wang C. -H., Chen, N., 2015. *A GIS-based spatial statistical approach to modeling job accessibility by transportation mode: case study of Columbus, Ohio*. Journal of Transport Geography, vol. 45, pp. 1–11.
- Warakomska K., 1970. *Ekwidystanta od dróg i od przystanków autobusowych a rozmieszczenie ludności w województwie lubelskim*, Lublin: UMCS.
- Widener M. J., Farber S., Neutens T., Horner M., 2015. *Spatiotemporal accessibility to supermarkets using public transit: an interaction potential approach in Cincinnati, Ohio*. Journal of Transport Geography, vol. 42, pp. 72–83.

