THE USE OF MODERN INFORMATION TECHNOLOGY IN RESEARCH ON TRANSPORT ACCESSIBILITY

Summary. Transport accessibility can be analyzed using a number of different methods. The problem with each of them is the difficulty of obtaining data to measure this phenomenon. The focus of this article and its main goal are to present methods and tools for gathering data on road traffic; thanks to modern information technology, it is possible to collect real data without the need for large-scale and highly capital-intensive measurements.

The application of modern information technology (IT) presented in the article, such as computer programs and applications like Google Maps Traffic Overlay and TomTom Live Traffic, enable research to be conducted on a scale that has thus far been unattainable, and allows information to be collected on such criteria as traffic volume, flow, average traffic speed, and actual journey time. Such innovative means of gathering data on automobile traffic open up new perspectives for assessing transport accessibility in terms of automobile traffic by providing high-quality data that meet the requirements for use in primary research.

ZASTOSOWANIE NOWOCZESNYCH TECHNOLOGII INFORMACYJNYCH W BADANIU DOSTĘPNOŚCI TRANSPORTOWEJ

Streszczenie. Dostępność transportowa może być analizowana z zastosowaniem różnych metod. Podstawowym problemem wszystkich wyróżnianych metod jest brak danych, które pozwalają na pomiar tego zjawiska. Istotę niniejszego artykułu i jednocześnie jego cel główny stanowi zaprezentowanie autorowej metody i narzędzi gromadzenia danych na temat ruchu drogowego, które dzięki zastosowaniu przez autorów nowoczesnych technologii informacyjnych umożliwiają gromadzenie rzeczywistych danych bez potrzeby prowadzenia zakrojonych na dużą skalę i wielce kapitałochłonnych pomiarów.

Zastosowanie nowoczesnych technologii informacyjnych zaprezentowane w niniejszym artykule w postaci autorskiego programu komputerowego oraz aplikacji Google Maps – Traffic Overlay pozwala na rzeczywisty pomiar dostępności transportowej w skali, która dotychczas nie była możliwa. Podejście to pozwala na wykorzystanie danych uzyskiwanych w czasie rzeczywistym, takich jak: wielkość i natężenie ruchu, średnia prędkość przemieszczania oraz czas podróży.
1. INTRODUCTION

A major research problem today in the field of transportation geography is finding means of obtaining data that would enable in-depth research to be conducted on transport accessibility based on actual automobile traffic. The only reliable source of information currently available for analysis of road load is General Traffic Measurements, which are conducted on regional and national roads every five years (the latest two took place in 2005 and 2010), and which allow total vehicle traffic on a given road can be determined. The issue remains, however, of how to measure the volume of traffic on individual road sections. In addition, General Traffic Measurements are only carried out on major regional and national roads that comprise important elements of transport networks. They enable one to analyze changes in road load, but not to define daily or weekly fluctuations in traffic intensity [13].

Analyses of transport accessibility (in terms of automobile traffic) enable us not only to diagnose the movement of goods and people, but also help to clarify a number of spatial socio-economic problems, and – perhaps most importantly – encourage the adoption of appropriate transport policies. For this reason, such analyses have proven extremely important. The conclusions drawn from accessibility analysis can be useful for decision makers and planners on various levels of management, from the local to national echelons [2].

The impact of transport accessibility on specific areas and locations, and transport strategies formulated on their basis are often evaluated by means of methods that decision-makers and scientists studying accessibility and mobility can easily be put into operation and whose results can be easily interpreted, such as travel speed, which, however, generally do not meet the criteria for use in theoretical research.

More advanced tools for measuring accessibility increase the complexity and effort involved in calculating and interpreting the results. Current practices in this area can be significantly improved by introducing more sophisticated means for operationalizing accessibility based on modern computer systems, using mobility data transmitted in real time [7].

Research on the development of more sophisticated means for measuring accessibility and mobility point to the need to consider the effects on accessibility of spatial and time constraints on individual units and feedback mechanisms between traffic on national and provincial road networks. This article responds to the ongoing need for theoretical and empirical research on the relationship between accessibility, the number of vehicles, and the speed of their movement.

Achieving good transport accessibility and evenly distributed transport services is one of the main objectives of transport policy. The method presented in this paper is a response to several constraints to measuring accessibility with other existing models used in planning practices. It offers an alternative and practical tool to help planners and decision-makers analyse the strengths and weaknesses of the transport infrastructure.

Transport accessibility can be analyzed by means of a number of different methods. Some authors list as many as a dozen or more [5]. Most authors agree that there are a few highly valid methods. The most widely used approaches include:

- Accessibility understood as the infrastructure equipment available to a given population or on given land units (the traditional approach);
- Time accessibility referred to cumulatively, i.e. according to the time needed to reach a given destination, which may be illustrated by isochronic maps;
- Accessibility illustrated by means of a possible return journey from a given destination within a given day, taking into consideration the infrastructure equipment and public transport;
- Potential accessibility expressed by means of a synthetic indicator measuring the total of distance between all elements (regions, sites) of a set, taking into account travel time and the importance (attractiveness) of a given unit in the system (demographic, economic or other potential) [2].

The above-mentioned research approaches apply methods which use as sources indirect data that do not reflect the actual situation taking place on roads. The core of the project hereby proposed and its main objective are the development of an innovative method for gathering data on road traffic, and
using it to assess actual transport accessibility in automobile traffic in Polish cities and towns. In the following sections of this paper, the term ‘transport accessibility’ will be used exclusively in reference to automobile traffic. The creativity of the method being proposed lies in its application of new information technologies and widely accessible databases like Google Maps Traffic Overlay and TomTom Live Traffic, thanks to which it is possible to monitor car traffic continually and in real time. These programs use signals sent by various types of equipment that are placed in vehicles moving along roads (i.e. mobile phones, GPS, some types of portable computers).

The database is also enriched with information received from taxi corporations and specially designed appliances placed on roads. The scope of this information is naturally not exhaustive in terms of all road users, but the universality of the above-mentioned equipment and annual increase in the number of users in Poland enable us to consider such a source of information to be highly reliable. According to studies conducted in 2012 (Interactive Advertising Bureau Poland, quote. at www.socialpress.pl), more than 6 mln Poles had a mobile phone with positioning function, and it was estimated that this number will increase to 8-9 mln by the end of 2016. If we add all road users’ equipment that sends information about the users’ positions (i.e. GPS), the suggested approach proves valid, especially in the long term, as the use of such equipment in the future will be as common as the use of mobile phones now.

Using databases such as Google Maps Traffic Overlay and TomTom Live Traffic enables one to obtain data from any period of time, any road, any road section and any number of roads. The only limit is computer performance.

Studying transport accessibility has been of interest for a long time, and in this sense, the research hereby proposed is a continuation of search on effective and ergonomic ways of measuring this phenomenon. At the same time, in comparison to existing methods, the methodological approach hereby presented is innovative on a wide scale. Creating a method for measuring actual transport accessibility will have a considerable impact on the approaches currently used in transport geography to measure this phenomenon, and will simultaneously indicate the possibilities for applying modern information technology in this field of study.

The suggested method represents a fundamental change in the approach taken towards the study of transport accessibility. By applying modern technologies and using widely available data (Google Maps Traffic Overlay and TomTom Live Traffic), the method allows for a precise and current description of phenomena such as automobile traffic volume, flow and average automobile traffic speed without the need to conduct expensive direct studies. Appropriately applied, the method enables researchers to go beyond the only method applied today for obtaining information on automobile traffic intensity, that is, the General Traffic Measurements conducted by the General Directorate for National Roads and Motorways (GDNRM). The proposed approach is not only cheaper, but it also enables the phenomenon to be monitored continually.

The new approach also allows for assessment of accessibility on the level of a single settlement (city, town or countryside); it is possible, among other things, to assess the influence of a narrow crossroads on the accessibility of a given cities/town or region. Additionally, the method could be used to conduct multidimensional analysis of actual transport accessibility of Polish cities/towns, which so far could not be analyzed because of a lack of data that would allow research to be conducted on such a large scale. The results and applied methodology could also provide a complex diagnostic tool for assessing the functioning of the transport system and evaluating infrastructural investments in both ex ante and ex post analyses. The scope of potential research is extremely wide and depends on the subject chosen for study. Limits result exclusively from the size of the available database and the optimization of equipment (i.e. the computer), both of which are used to conduct the research.

The data obtained by the method described herein can help improve the quality of representations of the spatial effects of interactions on the demand for transport in model studies, which are often based on generic curves that have little grounding in empirical data [10].

In terms of applicability, the research results provide diagnostic material for institutions dealing with the design and realisation of transport infrastructure, as well as for regional governments at all levels responsible for drafting spatial policy and strategic documents on i.e. the functioning of transport systems.
Accessibility is a term commonly used to assess the impact of transport strategies, but also in relation to land use spatial policy. It is usually assessed using separate models for transport and spatial analysis [12]. The method presented in this paper addresses both issues simultaneously. The postulated method also includes areas that can be further developed. The first is the determination of adaptive accessibility taking into account the factor of competition at the local level (e.g. workplaces and employees). The second area is determining the optimal scenario for implementation of policy measures using indicators for capacity and adaptive indicators of accessibility.

If the postulated method is used on a local scale, it can be used for planning sustainable accessibility in urban areas [6]. An example of another tool of this type is the SNAPTA model, which has been used in a pilot study in Edinburgh to determine what changes are necessary to improve the functioning of the public transport infrastructure in order to increase the accessibility of public transportation to the six types of activity [8].

Information and communication technologies allow users to access a virtual space without having to physically travel the distance and spend time travelling. Developments in technology, including those used in measuring accessibility, mean that distance recorded by means of different indicators is taking on new meaning. In physical space, accessibility is measured based on the spatial distance between sites or between possibilities and those seeking them [11]. From a virtual perspective, spatial distance loses its meaning. The study below makes it possible to incorporate the virtual component into existing available models. In general, the impact of the use of information and communication technologies in the field of accessibility cannot be overestimated.

2. EXAMINATION PROCEDURE

Today accessibility treated as physical distance is losing its validity, while the time needed to reach a given destination is gaining in importance. Among a number of methods that are applied for assessing the accessibility of a region or site, the one selected here is based on actual road load and average traffic speed on the road network.

As has already been mentioned, the initial data will constitute information received from sensors placed on roads, taxi corporations, government services, private distributors of databases and – most of all – anonymous users of the mobile version of Google Maps that has been developed for Android, IOS, Symbian, Windows Phone and other systems, as well as TomTom devices and software. The procedure for gaining information itself can be divided into steps. The first step is to define the site from which the source data is to be gathered. It is possible to use the services provided by the Google corporation, that is, Google Earth API, Google Maps API Traffic Overlay (developers.google.com, 2013) and TomTom Live Traffic. The information made available by these service providers include data on average traffic speed that are encoded by means of colours (tab. 1).

From a technical point of view, the information presented in Google Maps or TomTom applications are raster graphics images placed upon a given region after selecting an appropriate method from those available on Google Maps/Earth API/TomTom Live Traffic. Nonetheless, when the demo version of the program was being created, API elements did not include Poland (http://gmaps-samples.googlecode.com). As a result, gaining layers containing information on traffic in a direct way – using an open question – was not possible. As the program develops, it will be possible to analyze data not only in the form of raster graphic images, but also as databases.

To create the program use was made of object-oriented programming and a number of algorithms for analysing an image to allow for extraction of traffic speed. The first part of the information-gathering process is designing and coding a program providing data on road load.

The program has been written in JAVA and enables data to be read from selected elements of an image presented on a monitor, and modification of the flow to allow specific pixel colours to be read; a modified Thinning Algorithm [4] is used to decrease the number of pixels representing an individual reading, i.e. one vehicle. The technologies used are JAVA SE7, MySQL, Encog, and Google Earth.
Further development of the program’s functionality will concentrate on database analysis. The databases will be received directly from service providers and not via a raster graphics image.

### Table 1

<table>
<thead>
<tr>
<th>Colour</th>
<th>Description</th>
<th>Speed</th>
<th>Google interpretation</th>
<th>Author’s interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Google interpretation</td>
<td>Outside urban areas</td>
<td>Urban areas</td>
</tr>
<tr>
<td>green</td>
<td>More than 50 mi/h or 80 km/h</td>
<td>More than 80 km/h</td>
<td>- *</td>
<td>More than 60 km/h</td>
</tr>
<tr>
<td>yellow</td>
<td>25 - 50 mi/h or 40 - 80 km/h</td>
<td>40 – 80 km/h</td>
<td>- *</td>
<td>30 – 60 km/h</td>
</tr>
<tr>
<td>red</td>
<td>Up to 25 mi/h or up to 40 km/h</td>
<td>Up to 40 km/h</td>
<td>- *</td>
<td>15 – 30 km/h</td>
</tr>
<tr>
<td>black</td>
<td>Very slow - stop and go traffic</td>
<td>-</td>
<td>- *</td>
<td>0 – 15 km/h</td>
</tr>
<tr>
<td>grey</td>
<td>No data</td>
<td>No data</td>
<td>- *</td>
<td>No data</td>
</tr>
</tbody>
</table>

*Google API indicates differences between urban and rural areas, but does not indicate speed data.

The research question is answered based on the following procedure. Preparatory steps include primarily program calibration. This is needed because of the specifications of the computer equipment used. It is possible to conduct research on a number of computers working simultaneously, or on an appropriately effective individual device (a server) (fig. 1). The equipment needed at this stage can be rented from companies specializing in providing units for calculations.

In addition, calibration is necessary in order to define the region that will be analyzed on screen. The number of devices to be used is based on the extent to which the analyzed region will be covered and the required accuracy in pixel classification. The program gathers the data it will classify as Google Earth or TomTom speed measuring points. This is done on the basis of their colour and setting related to the defined road network.

The next two steps are: indicating four extreme x and y coordinates on the screen (\(x_{\text{max}}, x_{\text{min}}, y_{\text{max}}, y_{\text{min}}\)) that will set the field analyzed by the program, and assigning to the four coordinates digitals that refer to geographical width and length.

The preparatory procedure ends with defining the delay for starting the program, the number of analysis iterations in the selected field, and the length of the interval between iterations. For the sake of the project, the team set the number of milliseconds left before the hour 0:00 on the first day of the study as the delay for starting the program. An appropriate number of repetitions will also be set, with the restriction that the process will not be interrupted even if another process starts. The two processes will run simultaneously, as data analysis may take longer than a single cycle. The procedure is repeated by each device, and the moment it is completed, the initial state of the study is considered to be complete.
Upon starting the program, a new stage begins in which the pixels corresponding to the assumptions of the colour in the RGB colour model are selected. After selecting a field as a part of a layer, a tuple is created, which contains a symbol for the speed range, the number of the cycle, the number of the station, and the location of the pixel in the Cartesian coordinate system as reflected on the screen. After analyzing the whole image, the modified Thinning Algorithm is started, and those records that refer to one object are eliminated. The procedure contains a gradual elimination of the label that represents one object and lasts until a point consisting of many pixels no longer possesses the width and length of one unit. In terms of mistakes, it has been estimated that the algorithm makes errors in 15% of the individual labels selected (fig. 2).
The team realizes that this number is rather high. They have, nonetheless, decided to apply this solution to make sure that the data placed in the MySQL database are reduced. It is also possible to start the entry without decreasing the number of records. The results are similar because of the fact that average measurements are used in the analysis. A problem lies in interpreting the data in GIS, which in extreme cases would have to cope with a number of records collected in one day reaching $3 \cdot 10^6$. Thanks to the algorithm, this number has been reduced to $1.8 \cdot 10^6$. Application of this procedure is needed in order to limit the amount of data generated by the program. It also limits errors in measurement. This phase of the study is completed by recording data on every pixel in an individual line in the MySQL database. In order to store data in a safe but easily accessible way, it needs to be placed on a server. This device, like the other computers, can be purchased or rented from an external firm for the duration of the study and for a given amount of data.

After completing the required number of cycles, further steps include initiating a method for creating a network of squares. This method is used to calculate the number of occurrences of each type during the entire study. Its extent has been estimated on the basis of region size as obtained by the program, where one pixel in a preliminary examination responds to approximately 66 actual metres. Because of the duration of the procedure and the assumptions made, it was decided that the side of the square would be 50 pixels long. As a result, the created network consists of 5129 elements.

The next step is connected with starting SQL inquiries, the aim of which is to calculate the number of points in each selected square in terms of type (speed range) and to create a table for each of them. Later, a table is created from the data, with the results from each square being collected and adapted for interpretation by Geographic Information Systems (GIS). At the end of this phase of the research, it is necessary to export the data to a file (i.e. CSV) and import the data to GIS. For better visualization of the test results, it is possible to change the originally created polygon layer to a point layer using, for example, a Geospatial Modelling Environment.

After importing .shp files to GIS type software, spatial analysis of the data and its visualization will be conducted. Appropriate data averaging (methodology also used by the General Directorate for National Roads and Motorways) and appropriate classification will make it possible to estimate average speed, traffic volume and weighted density traffic on individual road sections. The length of the sections may reflect the ones applied by GDDKiA, but it is also possible to analyze sections of any length providing smaller or larger degrees of data aggregation. In addition, it is possible to analyze changes in traffic in reference to cities. Fluctuations connected with crossroads and road nodes may also be studied. When taking into consideration the costs of time, vehicle amortization, petrol (depending on speed) and possible payment for use of a highway, the costs of a journey on a given section and how they change due to traffic flow may be calculated. Thanks to the application of isolines (i.e. isochrone), the analysis is conducted in 3D. This enables definition of not only sites (i.e. towns and cities) but also regions with transport deficiencies.

In determining weighted traffic density, the following methodological assumptions have been adopted. The output speed adopted – distinguished in the study by exhibiting the largest number of repetitions – was not less than 80 km./h. In each of the primary fields, the weighted traffic density was reduced by the number of occurrences of each of the lower speeds multiplied by the appropriate weight expressing the strength of their impact on reducing the flow of traffic (a speed equal or greater to 80 km./h – weight 1; a speed of 40-80 km./h – weight 2; speeds in the range of 20-40 km./h – weight 3; speeds equal to or less than 20 km./h – weight 4).

3. PRELIMINARY RESULTS

The research project presented herein was preceded by a preliminary study conducted using a trial version of the program, performed on 21 December 2012. The network of national roads in the Łódź region was analyzed in terms of road load, traffic flow and average traffic speed. After comparing the transport network with the region’s settlement grid, it was possible to assess the transport accessibility of the towns/cities to which the roads lead. The observed vehicle flows corresponded to the results of the General Traffic Measurement done in 2010. The research differed, of course, in the raw number of
observed vehicles, which resulted from the different methodologies used in counting them, but the differences in road load on each section were the same.

From a dynamic perspective (daily or weekly), the research allowed us also to establish the fluctuations in traffic flows going through towns/cities in the region and, what follows, provided us with knowledge on how many passengers actually reached towns/cities (in contrast to assessments based on models that take into account technical road features or population density). Analyses of the results of the preliminary study, which applied the above-mentioned assumptions, indicated that in terms of cities, the highest transport accessibility was assigned to Łódź, and in terms of bigger towns to Łowicz, Piotrków Trybunalski, Radomsko, Skierniewice, Tomaszów Mazowiecki, Wieluń, and Zgierz, and for smaller towns to Kamieńsk, Rawa Mazowiecka, and Stryków. Taking into account the size of the city and towns seems reasonable, as differences in the load of roads leading to them are significant; nevertheless, the differences become smaller if we take into consideration their populations.

From the perspective of the research conducted, it is worth noticing that average traffic speed on national roads in the Łódź region is higher than 80 km/h, which reflects a high level of traffic flow on these roads. Slow vehicle movements are noticed only at crossroads and road nodes (fig. 3).

Fig. 3. Road traffic load in the Łódź region – test study
Rys. 3. Natężenie ruchu drogowego w regionie łódzkim – badanie testowe

The last criterion taken into consideration in analyzing accessibility to towns and cities in the preliminary studies was traffic flow. In terms of regions with the least traffic flow, they are comparable to the places with lower average traffic speed described above. Among these regions, however, there were some where accessibility was more limited than would be expected based on
The results of the preliminary study prove that the program created thus far is effective. They also show flaws in the program that result from the fact that the procedure has not yet been sufficiently optimized. However, these flaws will be eliminated within the frame of the project research. The preliminary study has also proven that our assumptions concerning the functionality of vehicle transportation in Łódzkie Voivodeship are valid. Because the preliminary study is not highly detailed (it was restricted to the national road network), the results should be treated as highly general. Nevertheless, they correspond to the results of the General Traffic Measurement method. Planned additional research will be much more specific, enabling us to present the analyzed phenomenon in more detail.

The influence of these research results on the scientific field is expected to be extremely significant. The scientific approaches applied thus far for analyzing transport accessibility will be verified by the use of a modern, efficient and ergonomic tool for gathering data to conduct research. The General Road Measurement used thus far may be ultimately replaced by the presented method, which would allow data needed for appropriate planning of the road infrastructure to be gathered more
inexpensively. The results would also provide diagnostic material needed for spatially drafting various strategic documents. From a dynamic perspective, the study will enable changes in traffic flow going through a given point to be established, and, as a result, will provide knowledge on how many passengers reach their destinations.

Fig. 5. Weighted traffic density on national and provincial roads in the Łódź region. Originally created polygon layer changed to a point layer

Rys. 5. Ważona gęstość natężenia ruchu na drogach krajowych i wojewódzkich w regionie łódzkim

Conducting research on the territory of Poland using the suggested method will enable us to assess the transport availability of towns and cities in terms of traffic flow and volume, as well as to measure average traffic speed on regional and national roads. The research conducted will also enable us to identify infrastructural transport demand-output mismatches in terms of both quantity and quality.

4. CONCLUSION

The primary incentive for initiating activities connected with the project was the possibility of creating a method that would enable one to obtain precise data needed for conducting further studies in the field of transport geography without the need to invest much money and to do so independently of other institutions and their research curriculum (i.e. General Directorate for National Roads and Motorways (GDNRM) and the study they organize every five years).
Today, automobile transport is the most developed branch of transport, and the traffic network remains the most highly-developed elements of the transport infrastructure. It is diverse in terms of its geometrical network, density, functionality, the accessibility of road nodes, and the way in which roads are connected with each other, settlements and neighbouring sites, and should be designed such that the needs of its users are met in the most effective way, creating at the same improved time traffic flow [3].

Taking into account the results of studies conducted using the method hereby proposed, planning documents could be drafted in a way that reflects actual traffic conditions. Such studies would also provide abundant diagnostic material on every territorial level. Any attempt to draw conclusions on the accessibility to individual regions or sites based on automobile movement as reflected in traffic volume on traffic routes connecting the settlements in a voivodship or country would be an exaggeration. Nonetheless, one may along with [9] assume that accessibility influences to a large extent the volume of traffic and as such constitutes an incentive for journeys. Therefore, it may be indirectly concluded that traffic volume on road sections leading to given destinations reflects at least the potential for reaching a given location by a defined number of vehicles. In the dynamic approach (i.e. a daily or weekly one), the study enables us to define the fluctuation in traffic going through a given site. Thanks to this it provides knowledge on the number of passengers that actually reach their destinations (in contrast to estimations based on models that take into account such things as the technical properties of roads or population density). Conducting research using the proposed method will enable one to conduct a total assessment of the actual transport accessibility of cities/towns from the perspective of traffic flow volume and average traffic speed of vehicles on regional and national roads. The research also enables one to identify infrastructural transport demand-output mismatches in terms of both quantity and quality.

What is more, based on this diagnostic data, guidelines for documents authorizing transport policy could be drafted. Thanks to the information on the weak spots in local and regional transport networks, it may be possible to choose which development paths are optimal and which regions need the most attention – that is which regions have such poor transport networks that they remain ‘excluded’ from transport.

The end effect will first constitute the optimization of an independent universal program used to gathering essential data and optimization of an algorithm for the applied procedure, both of which will be created for the needs of the research approach presented herein. Secondly, the result of the project will be a research method that will enable the assessment of actual transport accessibility that will offer highest – among the currently available methods – level of accuracy because it allows for an analysis of the phenomenon in reference to an individual settlement (city, town or countryside), and on various levels (traffic volume, flow and average traffic speed). The research will also provide abundant assessment material for creating documents of strategic importance on every territorial level.

References


Received 28.02.2014; accepted in revised form 20.08.2015