How to calculate real estate accessibility*

Piotr CICHOCIŃSKI, Poland

Key words: Geographic Information System (GIS), network analysis, attributes of real estate, accessibility

SUMMARY

Accessibility essentially describes an individual's ability to reach desired goods, services, activities and destinations. For many, buying or renting a property conveniently located within easy reach of work, schools or university is a major consideration. Properties with access to jobs, shopping and entertainment command higher prices and rents.
Accessibility of real estate is an attribute determining the possibility of reaching certain destination, driving one's own means of locomotion or using public transport. The term movement with own locomotion means refers to journey along the road network from real estate directly to some sort of central point, around which concentrates the life of inhabitants. Such analysis can be conducted independently for each real estate, by calculating best route to the centre, or zones of specified travel times around the centre may be determined and particular properties assigned to them. Regarding mass transport facilities, one has to move on foot to the nearest stop first to be able to ride further.
To perform any network analysis in a computer system, gathering and preparation of the data is required. OpenStreetMap – a community project aimed at creating editable and available without restriction map of the world. It can be recommended as an appropriate source of spatial data, particularly for network analysis.
Examination of the journey conditions was conducted basing on two data models: vector and raster. In the first case functions for geographical networks analyses were used. Vector analyses were limited to movement along elements of road network, unfortunately also for pedestrian traffic. As for rasters every variant of journey forced preparation of suitable data. Raster data also makes possible analyses regarding movement outside the road network, which can be applied in case of pedestrian traffic examination. After appropriate preparation of movement cost raster, it also appeared possible to integrate in one analysis getting to stop on foot and further traveling using public transportation means.
Research described in the paper confirmed the usefulness of tools for vector network analysis and raster cost analysis for determination of real estate accessibility. The only problem can be, as in many other cases, the availability of suitable data.

* This paper is connected with scientific works conducted in the Department of Geomatics, AGH University of Science and Technology

How to calculate real estate accessibility

Piotr CICHOCIŃSKI, Poland

1. INTRODUCTION

Accessibility essentially describes an individual's ability to reach desired goods, services, activities and destinations. For many, buying or renting a real estate conveniently located within easy reach of work, schools or university is a major consideration. Properties with access to jobs, shopping and entertainment command higher prices and rents. Accessibility of real estate is an attribute determining the possibility of reaching certain destination, driving one's own means of locomotion or using public transport.

The term movement with own locomotion means refers to journey along the road network from real estate directly to some sort of central point, around which concentrates the life of inhabitants. The assessment of such journey is influenced by two factors: the "distance" of real estate from a central point and the quality of the route that this access takes place. The word distance was intentionally enclosed in quotation marks here because it is not measured as the crow flies, but it corresponds to a distance one has to cross, moving roads. By contrast, the quality of travel is affected by difficulties one may encounter along the way. Drivers moving in urban areas are well aware of traffic jams formed during peak hours. The reliable analysis should also take such factors into account. Ultimately, these two elements determine the time of arrival, which can be considered as a measure of accessibility.

Getting to the center using public transport, at least in part covering the journey, is dependent on similar factors. The main difference, however, arises from the fact that first one needs to reach on foot to the selected station, and only from there one can leave using means of mass transport. Therefore, for the property owner, the distance to the nearest stop and wait time for the arrival of the next bus, tram or subway are also significant. Such a division into two, somewhat independent stages greatly makes modeling difficult. Furthermore, in case of pedestrian traffic it is not easy to determine which way it is carried out. In areas with dense buildings and fenced parcels, pedestrians move along the roads and streets. However, on areas not built–up or with sparse buildings, pedestrian traffic is usually possible in the whole area between the buildings. Therefore it is important to prepare the appropriate data, taking into account this diversity.

The above analysis can be performed independently for each real estate by calculating the best route to the center, or zones can be constructed around the center of specific travel times and individual properties can be assigned to them. The latter approach may be particularly useful in case of mass testing. The study can be conducted based on two types of data models: vector and raster.

Vector data in this case is a collection of interconnected objects, forming a road network: lines (edges) – representing the center lines of roads and streets, and points (nodes) – in a smaller scale representing cities, in the greater scales corresponding to street crossings (Curtin 2007). Along the edges the analyzed traffic is carried out. The nodes occur at the junctions of two or more edges and allow the flow between different edges. For the network to function as an accurate model of the real world, the additional information must be linked

to the edges, indicating whether the segment is passable in two or in one direction or impassable, and determining the cost of travel along individual segments of the network.
Depending on the context, this last parameter is called impedance or demand. The basic and easiest to obtain measure is in this case the length of the segment, which can be easily determined from the geometry of the object. However, more important attribute is the time needed to pass an element the network, which is the ratio of length and speed of movement. The value of this attribute is difficult to obtain, because one needs to know the speed. Additionally, to make the analysis closer to the reality the information about the cost of turns between the different sections of streets meeting at the intersection can be assigned to the nodes.
The second model used is the raster model. Selected part of space is divided into square grid elements, called pixels. Up to one value can be assigned to each pixel. In the case of using the raster model to analyze the movement this value is the impedance that must be overcome to get to the next pixel. Unlike for vector data all pixels are the same size only time can be used as impedance.
Various network analysis functions, implemented in geographic information systems (GIS) software, exploring the possibility of movement along individual components of the network operate on the vector data. The most widely used and most common one is the function of finding the optimal route between two points. In this case Service Area function, available in the ArcGIS software was used. The service area is the set of all lines (or polygon covering these lines), forming paths beginning at a center point, where the total impedance measured from the center point along every individual path will be no larger than a given value. An important drawback of the vector analysis is its limitation to travel only on the road network elements, unfortunately also for pedestrians.
In case of rasters Cost Distance function was used, which for each pixel determines the accumulated cost of reaching the nearest center. One advantage of this function is the ability to take into account the center of an elongated shape (for example located along the main road or street, which is often seen in rural areas). The second benefit is the ability to analyze the traffic that occurs off–road.

2. DATA PREPARATION

To perform any network analysis in a computer system, gathering and preparation of the data is required. Because of the popularity of GPS receivers and their wide application in car navigation, there are many companies offering network data sets. However, the acquisition of such data is associated with considerable costs. Therefore, an interesting alternative could be OpenStreetMap – a community project aimed at creating editable and available without restriction map of the world (Haklay, Weber 2008).
Such a map is created based on data from handheld GPS receivers, aerial photographs and other available data sources, as well as sketches made in the field. OpenStreetMap project was inspired by the idea of Wikipedia: map window in a web browser is provided with the “Edit” tab (fig. 2), and all changes are recorded. Registered users can add and modify data using the available editing tools. They can upload GPS tracks and edit vector data using applications available for this purpose. Collected information is stored in one central database, which can be downloaded in the form of a map image presented using the selected symbology (fig. 2) or

vector data. The project was created because most of the maps, which are generally considered free, actually have legal or technical restrictions on their use. There are requirements imposed on data added to OpenStreetMap: they have to be correct, verifiable and not being subject to copyright or the person entering data has to have full rights to them.

OpenStreetMap resources are made available under the Creative Commons Attribution–ShareAlike 2.0, which states that they may be copied, distributed, displayed and used for creation of derivative works provided that they are attributed in the manner specified by the author or licensor and the resulting work is distributed only under the same or similar license to this one.

Because this database is built by volunteers, no plans are formulated for its systematic development. Addition of new data depends on the willingness of individuals to perform the appropriate measurements on the ground or to vectorize available aerial photos. There are however cases, that companies or institutions having various data in their resources decide to give them free to OpenStreetMap community. Globally significant are permissions given by Yahoo and Microsoft for the unlimited use of the data (especially aerial photographs) provided by their mapping portals.

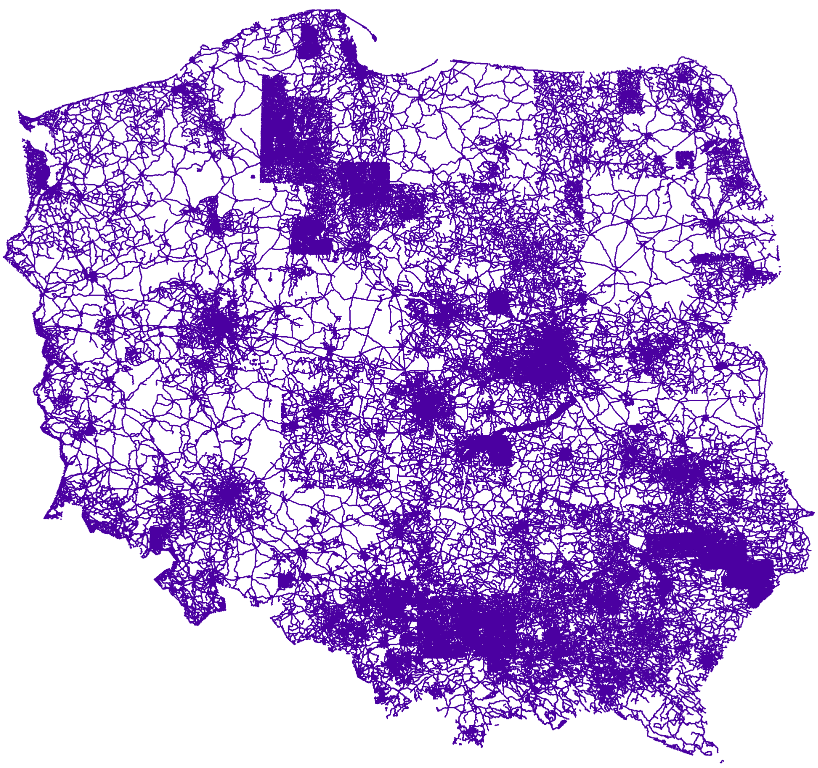

Fig. 1. Availability of information about the roads and streets network in OpenStreetMap database for Poland

The only significant drawback of this data set, resulting from the method of its creation, is the lack of central data quality control. Out of considered for geographical data quality aspects

(Cichociński 2010), in this case it is necessary to mention: accuracy, completeness, geometrical consistency and accuracy of the thematic classification. The variable streets and roads network completeness for Poland can be seen on the sketch presented in figure 1.

While the first two problems in practice require time–consuming comparison with actual course of roads and streets, the geometrical consistency may be checked using methods of internal evaluation. A topology used to describe the spatial relationships between geographical features may help to solve this problem (Cichociński 2008). In the case of such feature collection like road network, to check the geometric consistency it is enough to use only the two topological rules: “lines can not intersect and overlap” and “lines must only touch at their endpoints”. When detected these errors can be corrected using both automatic methods (for intersections) as well as semi–automatic or manual methods (in case of dangling nodes).

Another difficulty associated with the use of OSM data is ambiguity of the feature classification. Traditionally, features in geographic information systems are divided into points, lines and polygons. In the case of network analysis lines form the basis, but only those representing roads on which the traffic is allowed. Features in the OpenStreetMap database are divided into groups (categories). These include: roads, water, transport, each of which is divided into subgroups. For example, the category of roads include: freeways, expressways, links, but also residential and pedestrian zones, and stairs, therefore features, on which traffic generally does not take place. In general the streets in the cities have names, so mostly this attribute can be used to select appropriate features, but it is necessary to remember to reject impassable ones, but also to add segments that do not have a name (such as links).

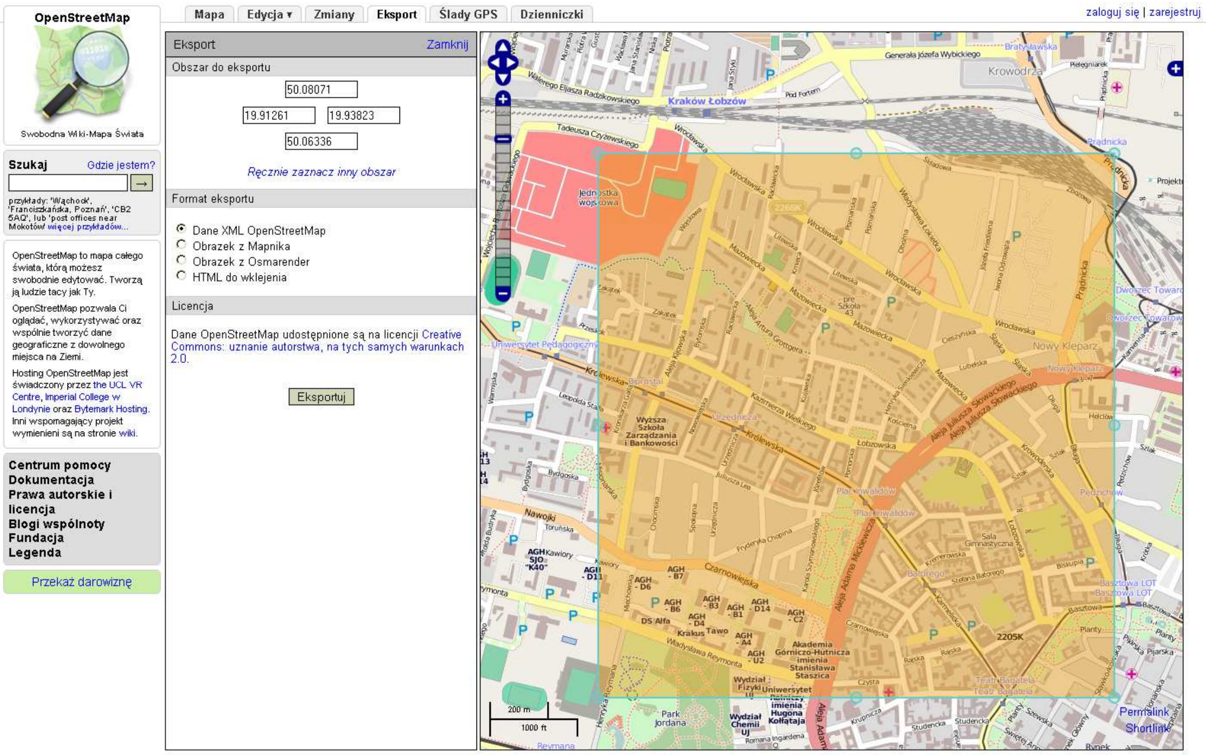

Fig. 2. OpenStreetMap window open in web browser, showing also data selection rectangle

Chosen fragments of the OpenStreetMap (OSM) database can be downloaded in two ways. The easiest way is to use “Export” tab, available in the map window (fig. 2). After determining an interesting range (by giving the coordinates or drawing the rectangle) an XML file containing features from the specified area is created. The necessity to possess the software allowing conversion of the OSM–specific format to one of the common GIS formats may be a potential problem in this case. An interesting proposal in this regard might be SpatiaLite spatial database management system.
SpatiaLite is developed by Alessandro Furieri, basing on another one–man project – SQLite. SQLite is actually a library implementing the self–contained, serverless and zero–configuration transactional database engine, managed through SQL commands. Because it operates on individual files, to some extent it can be compared to Microsoft Access, but its source code is in public domain. SQLite is the world's most widely used database. It can be found in such popular programs as Mozilla Firefox or Adobe Photoshop. There are unconfirmed reports on internet that Apple uses SQLite in iPhone and iPod.
SpatiaLite extends SQLite with the ability to store geometric characteristics of objects and to perform spatial queries. In practice, it consists of several programs, performing various specific tasks, executed from command line. One of them is a tool that converts a file saved in OpenStreetMap XML format to the spatial database table, immediately suitable for network analysis. Moreover, there are two more general tools, having a graphical user interface: spatialite_gui and spatialite_gis. The first one is a typical program for database management, allowing database creation and editing, shapefile import / export, SQL queries formulation and results displaying. The second program is a simple spatial data browser, allowing the visualization of query results in the form of simple maps.
Another possibility to obtain the OpenStreetMap data is to use one the following websites: http://download.geofabrik.de/osm or http://downloads.cloudmade.com. These sites provide data in shapefile format (ESRI 1998), which can easily be read by most GIS software. The region can be selected: the whole world, continent, country and, only for the first site, also the province. Partial thematic division is also made on this page (for example, roads and streets, necessary to perform network analysis, are saved into separate file).
The problem, which required special handling, was to combine in one study reaching to the bus stop on foot and subsequent traveling using public transport facilities. It seemed that it was only possible to carry out this variant using raster data, which had to be adequately prepared (Brzuchowska, Maśko–Osiadacz 1998). Both the road network, areas where pedestrian traffic is possible, and wait time at stops were modeled using the raster data.

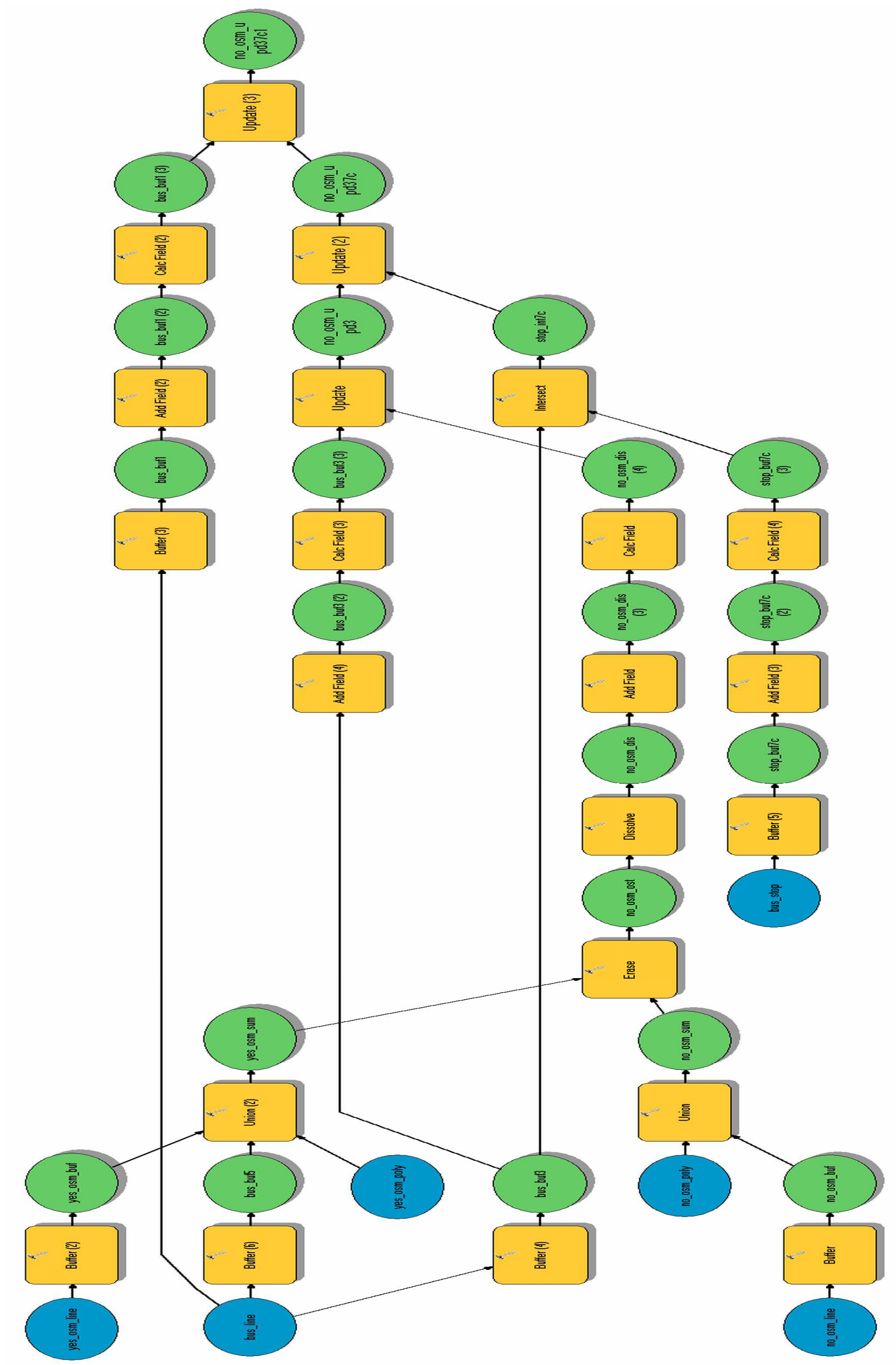

Fig. 3. Model for data preparation for raster analysis

For this purpose, analysis model shown in figure 3 was created in ArcGIS. Areas of roads, on which public transport lines run and where the traffic is the fastest were one of the results of its action (still saved in vector form) (fig. 4). They were surrounded by narrow areas described by attribute with a value of NoData, protecting against the possibility of leaving the road between stops. For the most part of the remaining space pedestrian traffic was assumed possible, with exception of barriers – places occupied by buildings and structures, rivers, streams and lakes, railway lines and restricted areas (such as military sites). Waiting time at the bus stop was recorded in a special way: buffers around the stops, greater than NoData area, linking pedestrian areas and routes, allowing "access" to stops were created. The attribute value assigned to them represented, in the form of impedance, the waiting time for a given stop. Afterwards, the resulting data set was converted to a raster with a pixel size of 1 meter.

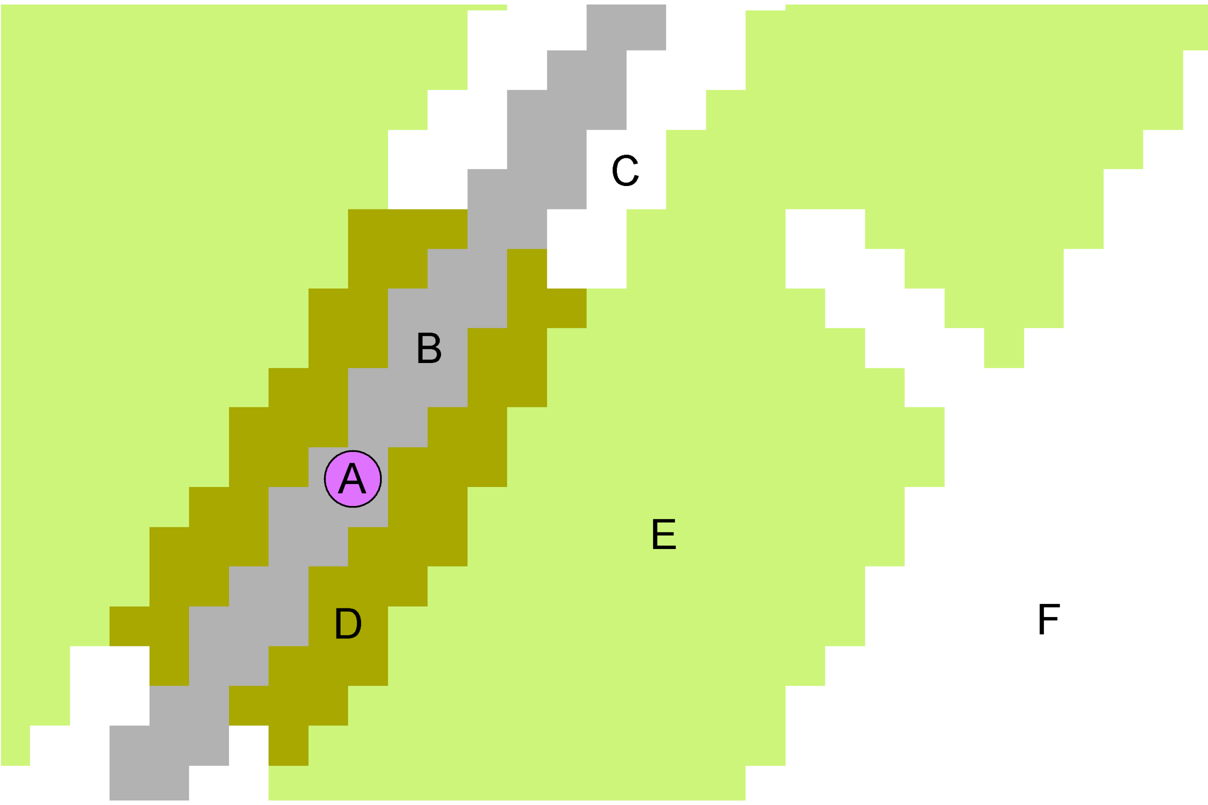

Fig. 4. Enlarged detail of data set prepared for raster analysis: A – stop, B – bus line, C – NoData area, D – "waiting time", E – pedestrian traffic area, F – barriers

3. RESULTS OF ACCESSIBILITY ANALYSIS

The following figures show results obtained from the proposed tools. The study was conducted in an area located in the northern part of Kraków, Poland. One of the important points to which residents commute from surrounding neighborhoods, is Nowy Kleparz. Several bus and tram lines reach there. The strict city center – the Old Town is situated not far from there. Appropriate data were obtained from OpenStreetMap (fig. 2). For the purposes of vector network analysis they were converted to shapefiles using SpatiaLite.

To avoid the necessity to analyze the availability for each property separately transport accessibility zone (at 5 minute intervals) were identified, to which different properties can be assigned depending on their location. Service Area function was used to perform this analysis for the case of one's own means of communication. Automation of this task also includes the possibility of making spatial joins for assigning properties to particular areas created around a central point.

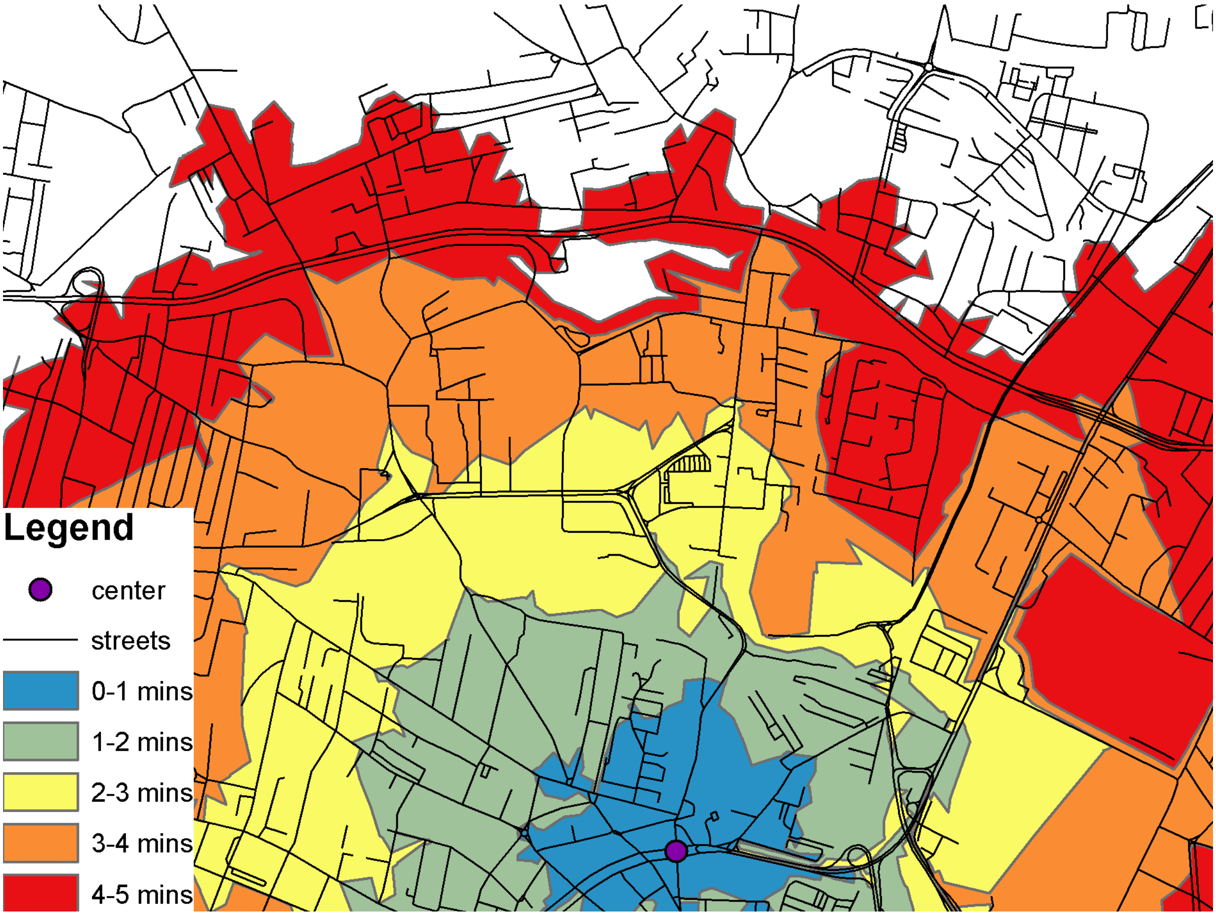

Fig. 5. Service areas calculated for the same speed for all streets

Two different results were obtained depending on the impedance adopted. In one case (fig. 5), the impedance was dependent on average speed fixed for all streets at 36 km/h, as determined in travel time studies conducted using the digital road atlas, while in the second case (fig. 6) different speeds were used for each category of roads, as proposed during the conversion process by SpatiaLite.

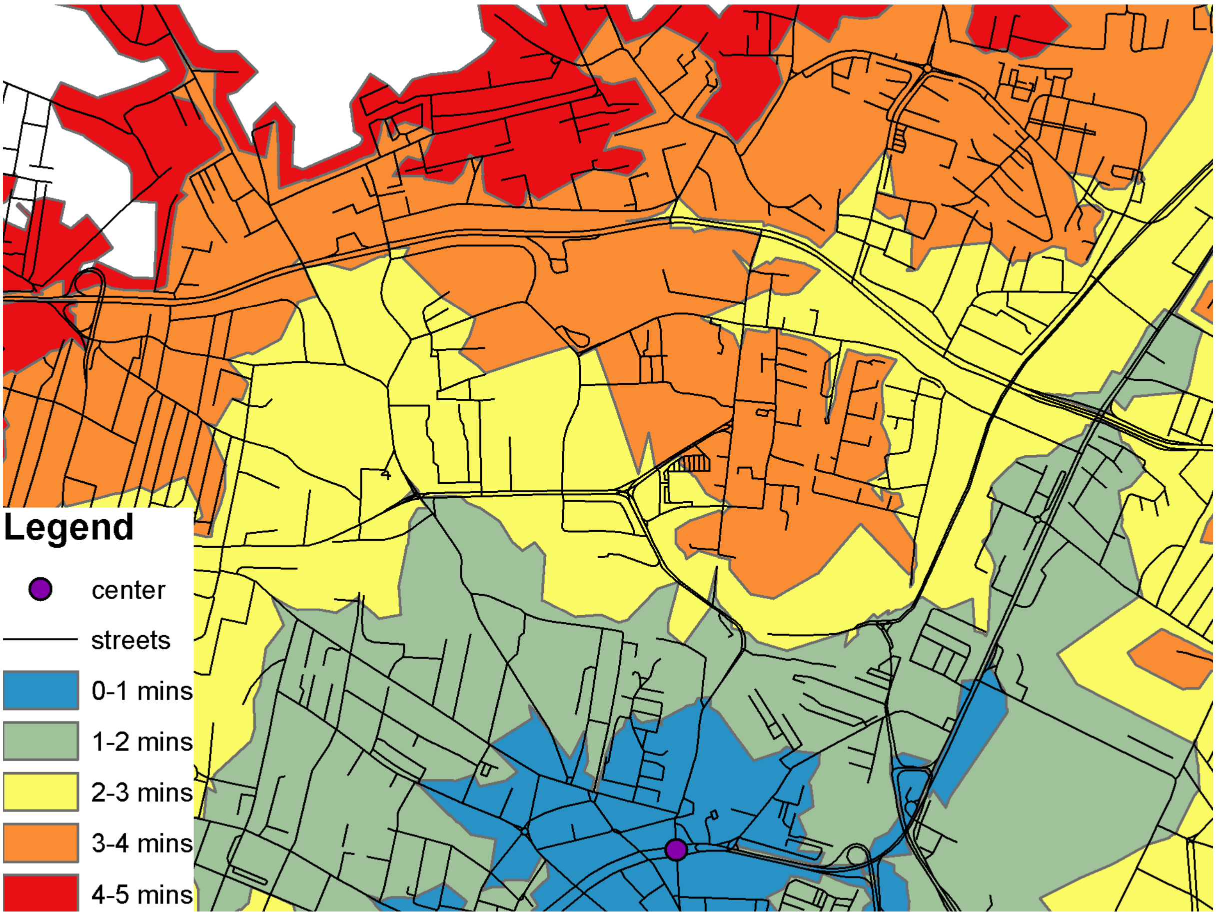

Fig. 6. Service areas calculated for different speeds for each road category

For the preparation of raster analysis, Quantum GIS software, equipped with the necessary plug–in, was used for making selection and saving in separate shapefiles the following groups of OpenStreetMap features:
– stops,
– barriers,
– streets, along which pedestrian movement is allowed.

Attribute value specifying the wait time for the vehicle, determined as the half of the time between successive departures (Transport for London 2010), was assigned to each stop.
These data sets were used as input for the model presented in figure 3, launched in ArcGIS. The resulting vector data set was converted to raster (fig. 4). The value 10, which corresponds to the velocity of 3.6 km/h (1 m/s) was assigned to pixels of size 1 meter, over which pedestrian traffic is possible. Speed of buses and trams was set at 18 km/h, which corresponds to pixel value 2.
Application of Cost Distance function, operating on the basis of so–prepared raster data allowed integrating in one analysis getting to stop on foot and further traveling using public transportation means. As a result of Cost Distance function the raster was obtained having pixel values corresponding to travel time from location represented by given pixel to the central point. Figure 7 presents this raster after reclassification made to introduce several time intervals of size 5 minutes.

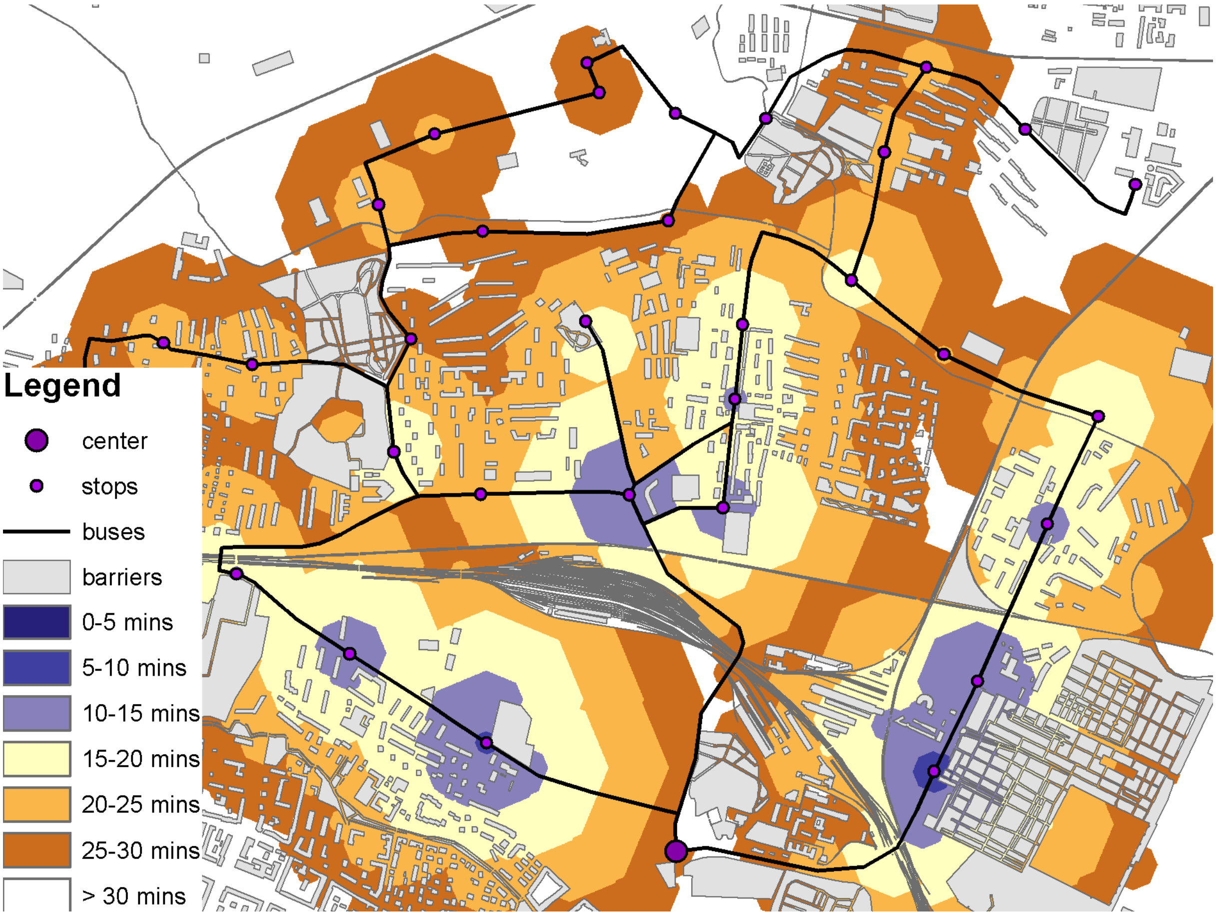

Fig. 7. The result of analysis integrating getting to stop on foot and further traveling using public transportation means

4. CONCLUSION

Network analysis performed on the vector data show the importance of proper attribute values describing the movement cost. Speeds proposed by SpatiaLite for each class of roads, and thus travel times appeared to be too "optimistic" compared to the everyday experiences of the author, who commutes one of the presented routes. It would be valuable to define and use actual traffic speeds, which can be obtained from GPS receivers (Cichociński 2011).
Limiting the analysis conducted on the basis of the vector data only to travel on the roads leads to the conclusion that the analysis based on raster data provide better opportunities for mapping various aspects of real estate availability. However, the raster form of the results is of little use. One must admit that although it is perfect for presentations, it is impossible to determine attributes of the real estate on that basis. For that reason, the resulting raster files after the appropriate classification of the pixel values must eventually be converted into a vector containing the zone boundaries, to enable assignment of particular properties to them.

Fig. 8. Enlarged analysis result for housing estate Żabiniec

Detailed analysis of the results obtained from raster analysis allows noticing certain imperfections of Kraków's transport network, as well as shows on what aspects of data preparation attention should be paid. For instance, housing estate Żabiniec (fig. 8), although located relatively close to the center, is characterized by difficult access by both public and individual transport means.

Although not entirely in accordance with reality, stops to be analyzed should be placed at street crossings. Otherwise some of the adjacent streets can prevent direct access to certain areas, what is clearly visible for example in figure 9. It turned out that many potential barriers to pedestrians outside the streets actually does not affect the result of the analysis, while elongated objects such as railway tracks or rivers, which have relatively few passages, proved to be very important. This can be seen especially in the case of the above mentioned housing estate Żabiniec, which is adjacent to the railway line. Also the problem of communication lines forming a loop was recognized. Introducing the entire route to the analysis may result in finding the best route having direction opposite to the traffic. For the future, it makes sense to include only that part of the route that leads from the end stop to the center.

The overall conclusion is, that particularly the analysis of raster data well fulfilled their role, especially that their made it possible to cover in one model both getting to stop on foot and further traveling using public transportation means. Of course, no analysis would be possible without having the relevant data. It seems that OpenStreetMap can be recommended as an appropriate source of spatial data, particularly for network analysis.

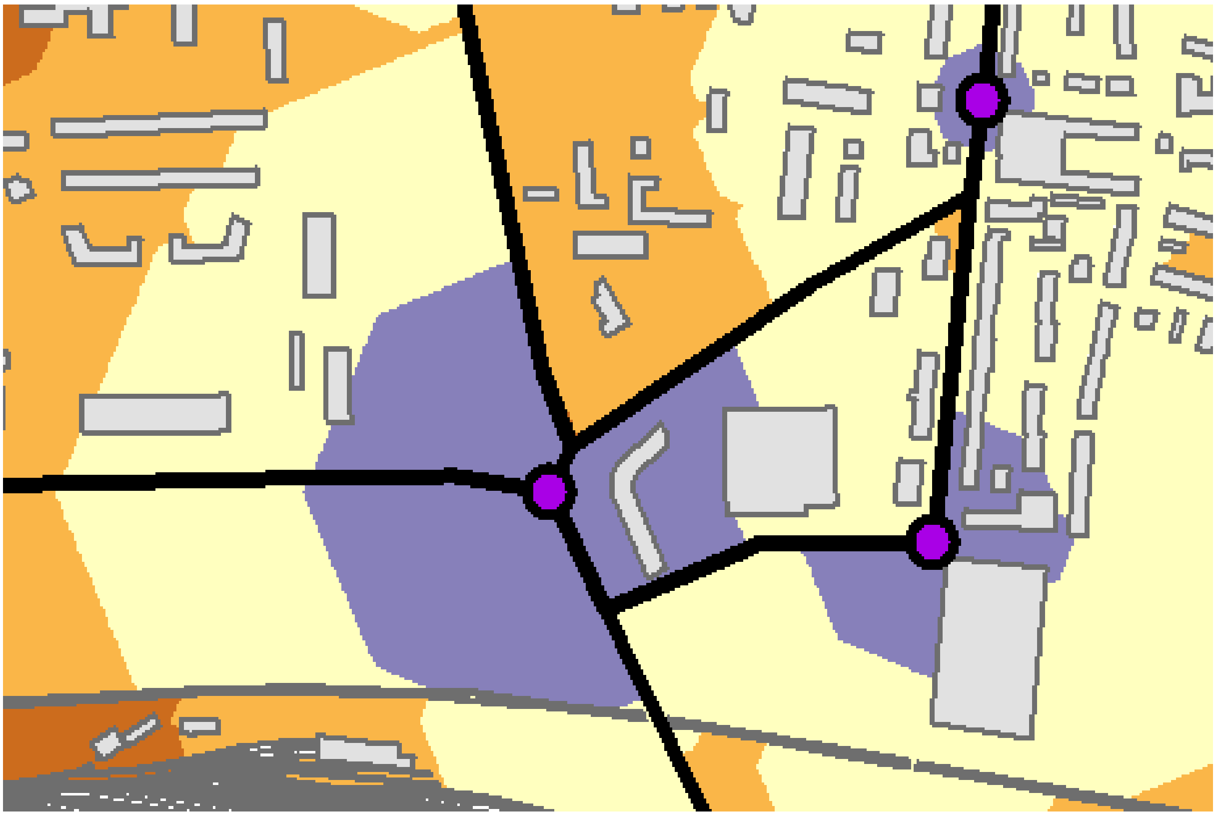

Fig. 9. An example of erroneous influence of street preventing direct access to certain area

REFERENCES

Brzuchowska J., Maśko–Osiadacz J., 1998, Zastosowanie struktury GRID do modelowania dostępności w ramach analizy warunków turystycznych regionu, Polskie Towarzystwo Informacji Przestrzennej, VIII Konferencja Naukowo–Techniczna Systemy Informacji Przestrzennej, Warszawa, 19–21 maja 1998 r.

Cichociński P., 2008, Application of advanced topological rules in the process of building geographical databases supporting the valuation of real estates, Proceedings of FIG Working Week 2008, 14–19 June 2008, Stockholm, Sweden.

Cichociński P., 2010, Metadane i jakość danych w systemach informacji geograficznej (Metadata and data quality in geographical information systems), III Krajowa Konferencja Naukowa „Technologie Przetwarzania Danych", Materiały konferencyjne, Poznań, 21–23 czerwca 2010 r., WNT, Warszawa.

Cichociński P., 2011, Zastosowanie techniki GPS oraz funkcji analiz sieciowych oprogramowania GIS do optymalizacji ruchu drogowego w miastach (Application of GPS technology and network analysis functions of GIS software for optimization of city traffic), Geomatyka I Inżynieria : kwartalnik naukowy Państwowej Wyższej Szkoły Techniczno–Ekonomicznej w Jarosławiu, 2011, nr 1, s. 13–25.

Curtin K.M., 2007, Network Analysis in Geographic Information Science: Review, Assessment, and Projections, Cartography and Geographic Information Science, Vol. 34, No. 2, 2007, pp. 103–111.

ESRI, 1998, ESRI Shapefile Technical Description, An ESRI White Paper, Environmental Systems Research Institute, Redlands.
Haklay M., Weber P., 2008, OpenStreetMap: User–Generated Street Maps, IEEE Pervasive Computing, October–December 2008, pp. 12–18.
Transport for London, 2010, Transport assessment best practice, Guidance document.

CONTACTS

Dr. Piotr Cichociński
AGH University of Science and Technology
Department of Geomatics
Al. Mickiewicza 30
30–059 Kraków
POLAND
Tel. +48 12 617 34 31
Fax + 48 12 617 45 88
Email: Piotr.Cichocinski@agh.edu.pl