Application of Advanced Topological Rules in the Process of Building Geographical Databases Supporting the Valuation of Real Estates

Piotr CICHOCIŃSKI, Poland

Key words: topological rules, conflation, real estate, mass appraisal, GIS.

SUMMARY

An essential attribute of real estate, influencing its cadastral value, is the location, and thereby the neighbourhood, both of other real estates, and also another objects. In order to carry out the valuation efficiently, spatial information on these objects should be collected in the suitable database.

The paper presents the usage of topological rules available in advanced GIS software to describe required mutual spatial relationships between objects coming from different datasets and to analyse occurring rule violations. Besides methods of detection of these inaccuracies, the methodology of their removal was proposed as well, using analytical and data processing GIS tools. Technology called conflation can also be used here. It refers to joining of at least two different data sets to create a new dataset, making suitable modifications of objects from both datasets. To carry out this operation on a broader scale it is required to use tools automating this process to the greatest degree, to reduce participation of the operator to interventions in a few non-typical cases.

Application of Advanced Topological Rules in the Process of Building Geographical Databases Supporting the Valuation of Real Estates

Piotr CICHOCIŃSKI, Poland

1. INTRODUCTION

For the valuation of real estate (and particularly the mass appraisal) to be efficient, spatial information on these real estates, as well as on other objects located nearby and influencing its cadastral value, should be collected in the suitable database (Cichociński, Parzych, 2005). It could come from GIS databases, including the land and buildings cadastre, the spatial registration of utility infrastructure, topographical maps with the digital terrain model, spatial development plans, as well as various environmental datasets. It can be supposed that the necessity of integration of these datasets will also appear during creation of national spatial data infrastructures.

However, before loading into one database, it seems necessary to perform mutual adjustments of these datasets. For example, former experiences of users working in Poland with various topographical databases showed the possibility of appearance of considerable inconsistencies between them of both geometrical and attribute character, due to different temporal periods and methods of data acquisition, and also the data sources. It can be supposed that the same problem will happen with the large scale data. It is at least partially confirmed by the analysis of parcel boundaries and spatial development plan zones coincidence, which were conducted by the author on the exemplary land cadastre precinct (fig. 1).
Necessary operations can be divided into two stages:

1. localisation of discrepancies,
2. removal of identified discrepancies.

2. TOPOLOGY

Completion of the first stage will be possible because of employment of topological rules – mechanisms of description of spatial relationships between neighbouring or nearby objects – available in advanced GIS packages.

Originally and in the wider meaning, the concept of topology referred to the branch of mathematics dealing with geometrical figure proprieties, which do not change under geometrical transformations (Magnuszewski, 1999). On the contrary, in the domain of Geographical Information Systems (GIS) topology is understood as a description of spatial relationships among neighbouring or lying nearby objects (Theobald, 2001). A pioneer in the field of topology application to reduction of the quantity of errors made in the process of large data sets collection was Unites States Census Bureau on the turn of sixties and seventieth of the 20th century. First topological rules based on the assumption, that objects lie on the plane and are represented by nodes (zero-dimensional), edges, called also arcs (one-dimensional) and polygons (two-dimensional). Arcs “must not intersect and must not overlap” but “must only touch other lines at endpoints represented by nodes”, forming non-overlapping polygons, filling the whole area. However the development of spatial data formats enforced taking into consideration mutual relationships between polygons and creation of the following rules: “must not overlap” and “must not have gaps” (fig. 2).

Figure 1. Discovered discrepancies between parcel boundaries and spatial development plan zones.

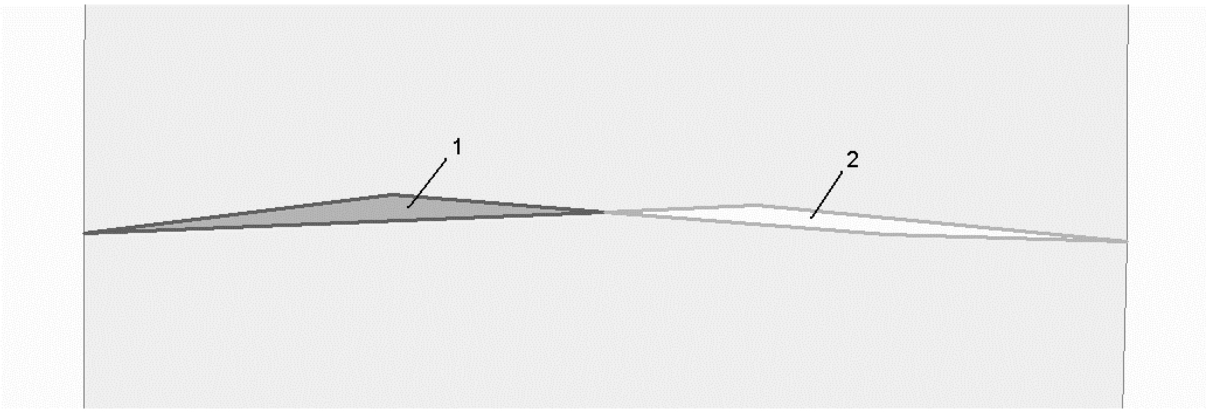

Figure 2. An example of symbolization of overlapping areas (1) and gaps between them (2).

Moreover the definition of topology was enhanced and now it includes the description of relationships between objects belonging to different feature classes. The list of available rules was considerably extended. Among other things it now contains the following statements: “boundaries of polygon features in feature class A must be covered by boundaries of polygon features in feature class B”, “objects of feature class A must be contained within polygons of feature class B”, “objects of feature class A must be covered with objects of feature class B”.
It significantly extended the functionality of this mechanism and enabled its application to the purposes presented below. It is however necessary to remember that detected violations of topological rules are not necessarily errors. Figure 3 perfectly illustrates that case, presenting buildings situated on two parcels. It sometimes happens in Poland (Eckes, 2002).

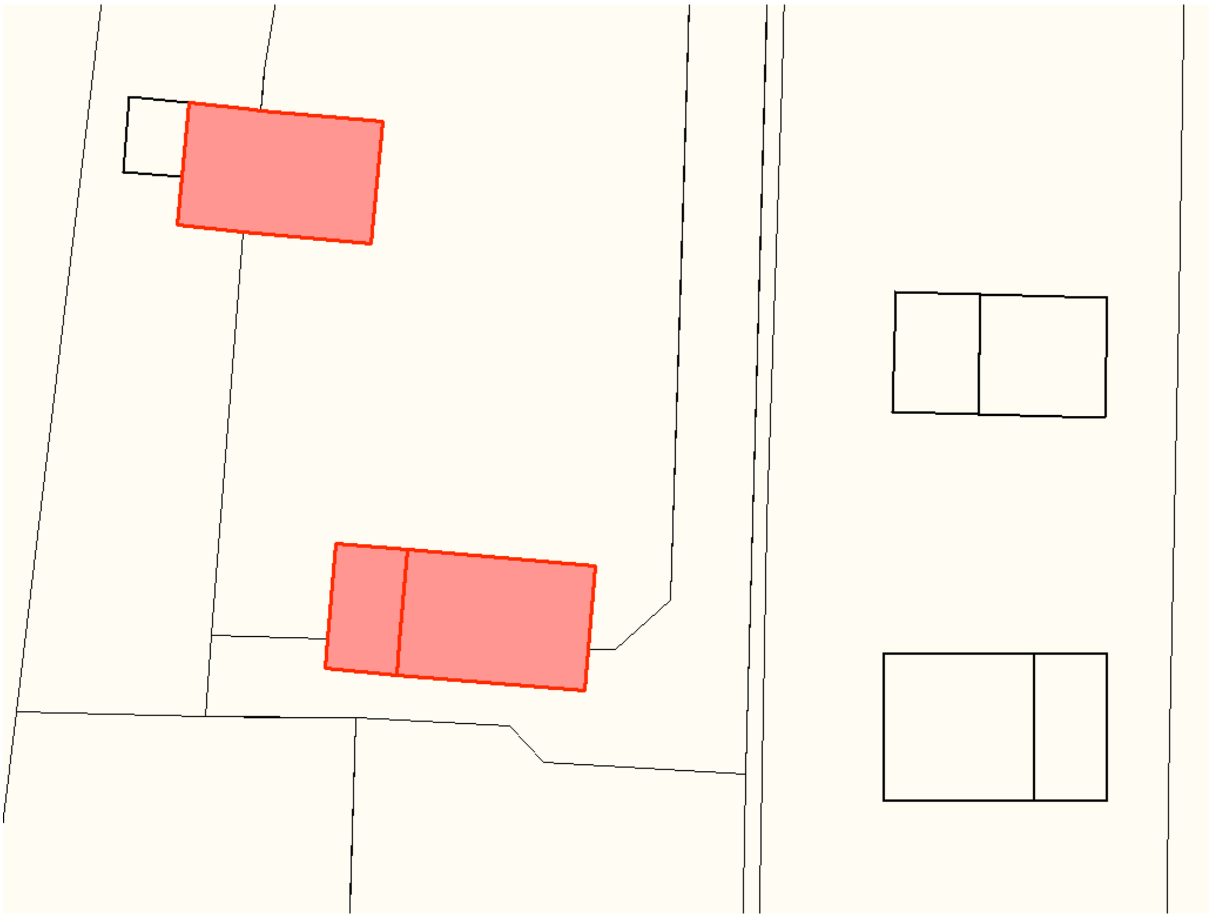

Figure 3. An example of topological rule “objects of feature class A (buildings) must be contained within polygons of feature class B (parcels)” validation result.

3. ERROR DETECTION

Through the analysis of possible cases, two groups of problems were distinguished and the method of solving them, using available tools, was proposed:

- ("horizontal") adjustment of adjoining objects (polygons) boundaries, originating from datasets having the same thematic content. An example of such case can be creation of continuous datasets by joining smaller fragments, such as map sheets or cadastral precincts. The solution of this problem is relatively simple. It consists in combining datasets into one, and then applying topological rules examining its internal correctness: "must not overlap" and "must not have gaps". The distinction of areas not satisfying these rules is the result of this operation (Fig. 2).
- mutual ("vertical") alignment of the position of objects belonging to different thematic layers, but having the same location (in other words the alignment of different partitions of the same fragment of the ground). An example of this type of operation can be alignment of chosen sections of parcel boundary and chosen sections of landuse boundary. Comparing to "horizontal" adjustment, the solution of this case is not univocal, because not all boundaries must coincide. The topological rule "boundaries of polygon features in feature class A must be covered by boundaries of polygon features in feature class B," is useless in this case, because it verifies the coincidence of all without exception boundaries, which by definition cannot be true. Therefore it is necessary to distinguish sections of boundary which should be coincident, from remaining ones, for which this coincidence is not required. It is possible in this case to use two criteria resulting from adopted assumption that locational differences do not exceed certain threshold values. The offset size and the angle between them are measures of such discrepancies.

4. REMOVAL OF DETECTED ERRORS

The best solution (in the second stage) of found discrepancies removal would be their analysis by the operator, and then manual correction using available editing tools. However, in the case of huge data sets or large quantity of found discrepancies, this method would be very time-consuming and would cause excessive costs of such operation. That is why the author decided to check, whether a tool making possible the automation of this process exists. The Java Conflation Suite (JCS) software is the result of this investigation. This is the toolset enabling topological validity of data sets verification and, to the certain extent, their automatic removal. It is built basing on technology called conflation, which consists of joining at least two different, neighbouring datasets, to create a new dataset. Therefore it is necessary to apply suitable modifications to objects from both datasets.

The software was created by the Canadian firm Vivid Solutions for the government of British Columbia province, looking for efficient, highly automated software for spatial data processing. An additional advantage, except the functionality, is that it belongs to the group of the free software, which makes possible not only cost-free usage, but also modification of the source code and adaptation to user requirements (Michalak, 2007).

However this is not an independent program, but rather suite of functions operating in the environment of Java Unified Mapping Platform (JUMP) – software for visualization, editing and analysing of spatial data, simultaneously allowing easy extension through special-purpose modules intended to solve various problems.

Java Conflation Suite offers two tools intended for this task: Boundary Match and Coverage Alignment. The first one, basing on the constancy of boundaries from one dataset, adjusts to them boundaries from the second dataset, located not farther than the established threshold value (fig. 4).

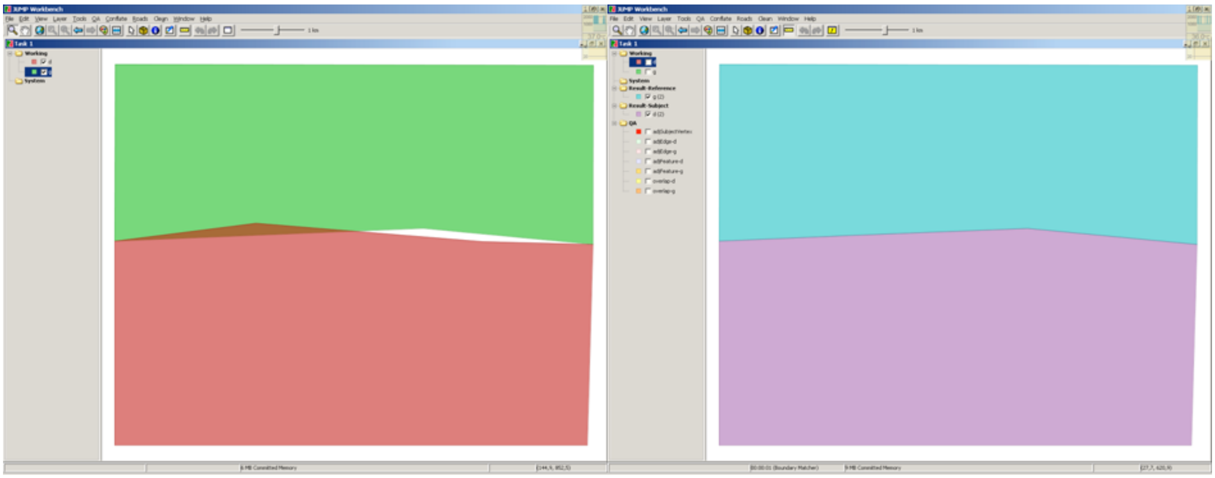

Figure 4. Input data and result of Boundary Match function.

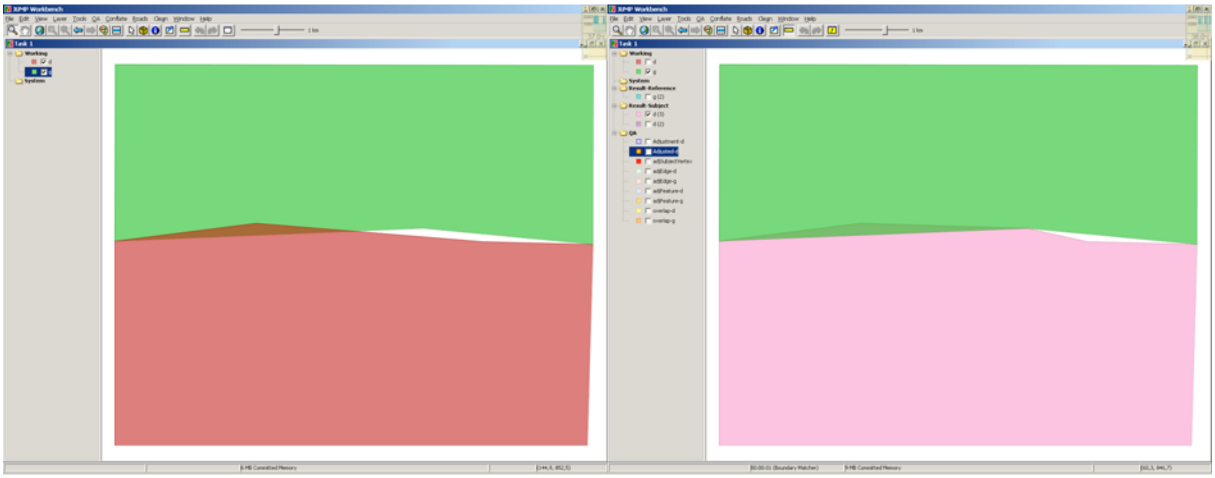

Figure 5. Input data and result of Coverage Alignment function after its first ru**n.**

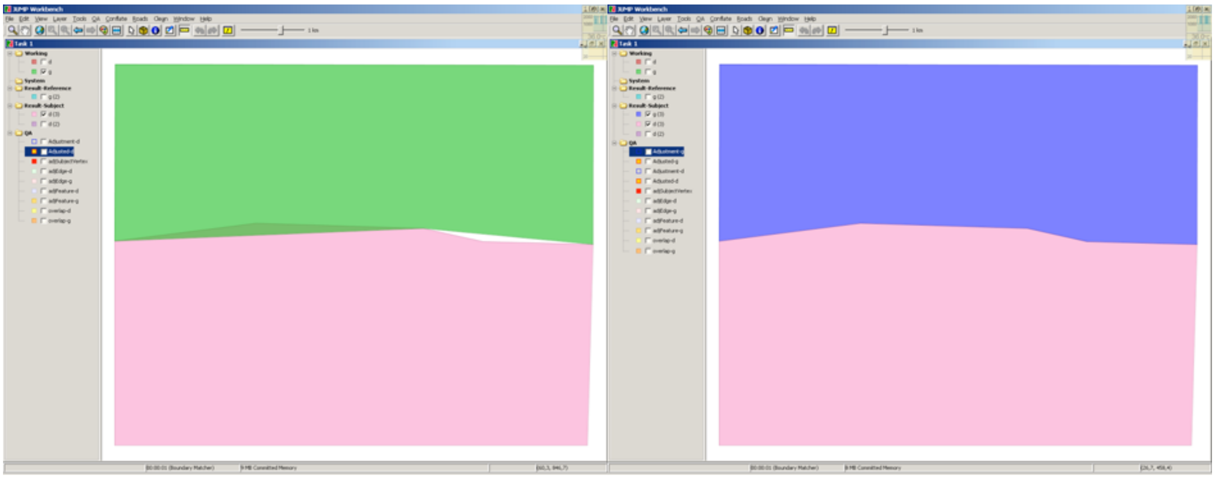

Figure 6. Input data and result of Coverage Alignment function after its second run.

On the contrary the function Coverage Alignment, taking into account two tolerance parameters described above (offset and angle between objects), after first run causes only partial amendment of errors, modifying the shape of objects coming from only one thematic layer (fig. 5). The second run is necessary, but this time taking into account second, still uncorrected dataset and adjusted version of the first dataset, what finally causes the removal of all errors (fig. 6).
The difference in results from these two commands is that the first one preserves boundaries of one chosen layer, while second causes the alteration of two layers.

5. ANOTHER APPLICATIONS OF JAVA CONFLATION SUITE

Java Conflation Suite has at least one more advantage: it can also be used for discrepancies detection. Function Find Misaligned Segments was designed for this purpose. The easier task is to find differing but matching outer boundary lines in compared datasets (fig. 7).

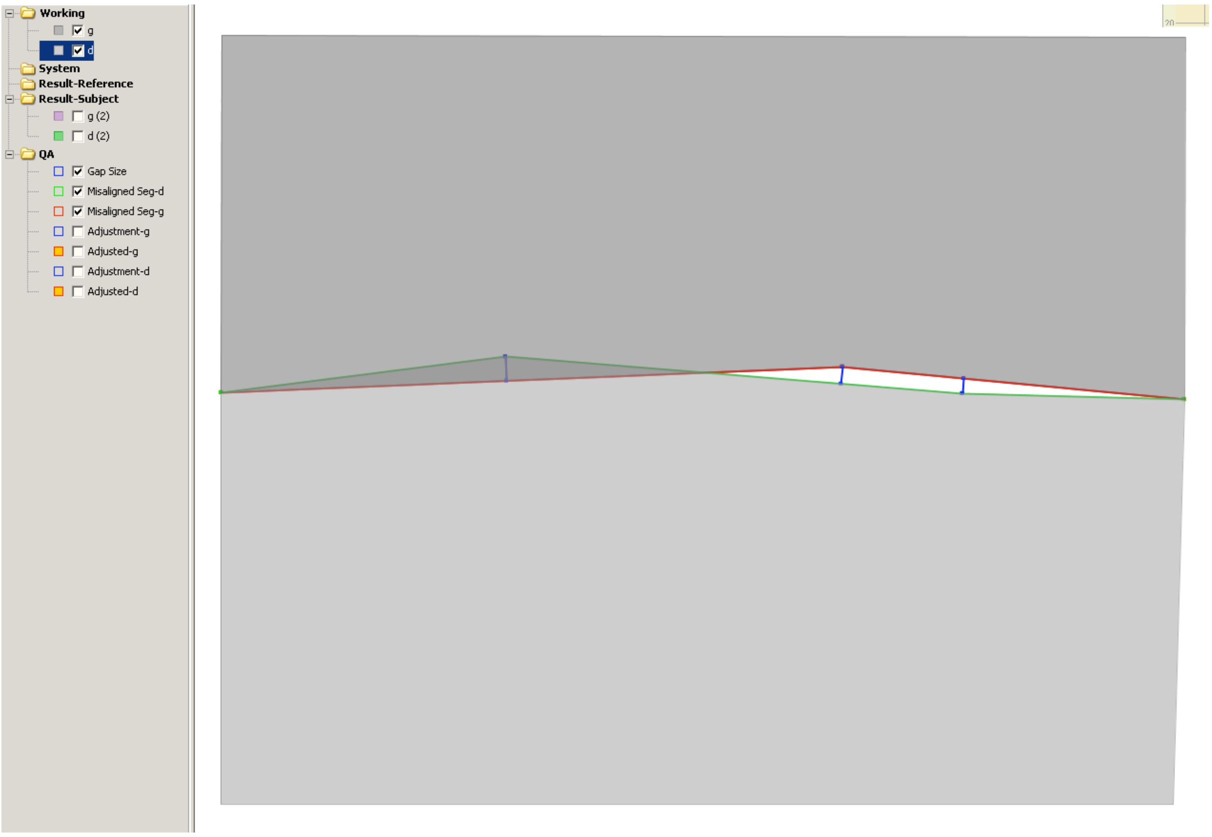

Figure 7. Corresponding, but differing outer boundary lines in compared datasets.

More complicated task is "vertical" adjustment. The function Find Misaligned Segments operates basing on objects comparison, with the regard to two above depicted parameters (offset and angle), whose values have to be smaller than the given thresholds. Sections of boundary lines from both compared datasets, lying not farther to each other than the value specified by the user and having angle between them smaller than given value (fig. 8) are results of this operation. Additionally the gap size "materialized" in the form of object in corresponding working thematic layer and simultaneously recorded as the attribute of this object allows evaluation of every error by the operator. Obviously, as it can be seen on figure 8, this function does not operate faultlessly, but just particularly in the case of shown discrepancies between local spatial development plan zones (the dashed line) and parcels (the continuous line), subtle intentions of planners can make correct operation of the automaton impossible.

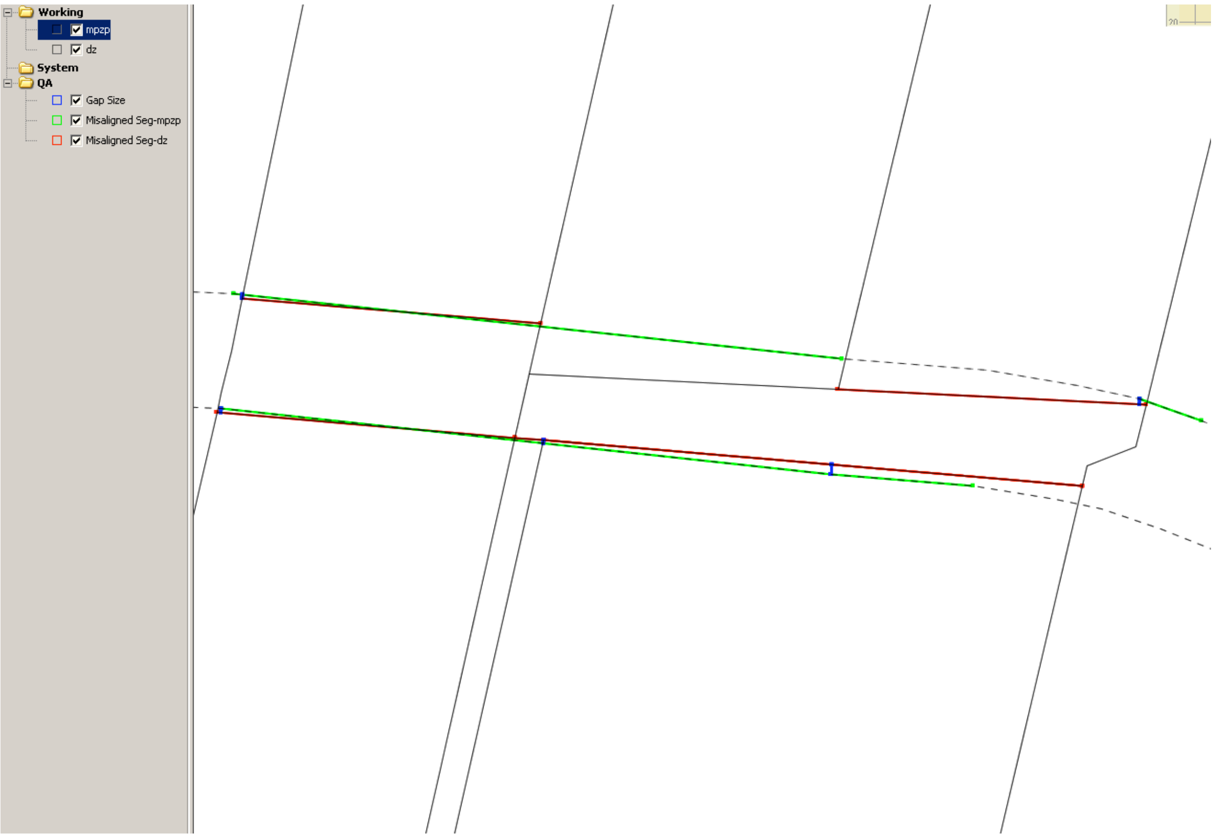

Figure 8. Results of Find Misaligned Segments function.

6. CONCLUSION

Summing up above deliberations it can be stated that research conducted by the author indicated that technical possibilities exist (in the form of suitable methods and tools) to build seamless databases for valuation of real estates – only suitable legal and administrative decisions are necessary. It is supposed that results of these works will be applicable not only to the valuation of real estates, but in every case, wherein the usage of many, potentially inconsistent sources of geographical data, is required.

REFERENCES

Cichociński P., Parzych P., 2005. Zastosowanie baz danych Systemów Informacji Geograficznej do wyceny nieruchomości. (Application of Geographic Information System databases for real estate valuation). IV Krajowa Konferencja Naukowa ”Infobazy 2005 – Bazy danych dla nauki”. Gdańsk, 25-27 września 2005 r.

Eckes K., 2002. Logika obrazu mapy jako podstawa do budowy systemu ekspertowego. (Logic of the map image as basis for building of an expert system). XII Konferencja Naukowo-Techniczna Systemy Informacji Przestrzennej. Warszawa, 17 maja 2002 r.

Michalak J., 2007. Otwarte oprogramowanie i otwarte dane w geomatyce. (Open software and open data in geomatics). Roczniki Geomatyki, tom V, zeszyt 2. Polskie Towarzystwo Informacji Przestrzennej, Warszawa.

Theobald D. M., 2001: Understanding Topology and Shapefiles. Arc User, April-June. Environmental Systems Research Institute, Redlands.

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 22 77
Email: Piotr.Cichocinski@agh.edu.pl