Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 447–457, 8 figs., 1 tab.

Semantic Data Integration: Data of Similar and Different Scales

Birgit Kieler & Monika Sester, Hannover; Haiqing Wang & Jie Jiang, Beijing, China

Keywords: Data Integration, Data Matching, Ontology

Summary: The integration of data of different origin leads to a lot of benefits: firstly, the properties of the individual data sets can be exchanged and used for mutual enrichment and benefit, secondly, the information from both sources can be "adjusted", leading to a more precise and reliable information. Such integration presumes that semantic relations between the data are known: the semantic correspondences help to enrich similar object classes and objects. If such semantic relations are not explicitly known, our approach is to exploit geometric correspondences of individual object instances stemming from different sources. From these geometric correspondences inferences concerning their semantic relations can be drawn. For these analyses, differences with respect to geometric and semantic granularity have to be taken into account. In this paper, we describe approaches to solve these problems.

Zusammenfassung: *Semantische Datenintegration: Daten ähnlicher und verschiedener Maßstäbe.* Die Integration von Daten unterschiedlicher Herkunft ist von großem Nutzen: einerseits können die Vorteile wie auch die Reichhaltigkeit der individuellen Datensätze ausgetauscht werden, um damit gemeinsam eine Verbesserung zu erzielen. Andererseits können die Informationen beider Quellen einander angepasst werden, wodurch eine höhere Genauigkeit der verfügbaren Daten entsteht. Solch eine Integration setzt voraus, dass die Beziehungen zwischen den Daten in einem bestimmten Ausmaß bekannt sind. Die semantischen Korrespondenzen helfen dabei, ähnliche Objekte miteinander zu vergleichen. Sind diese Relationen nicht explizit bekannt, nutzt unser Ansatz die geometrischen Korrespondenzen der individuellen Objektinstanzen der verschiedenen Datenquellen. Von diesen geometrischen Korrespondenzen können Rückschlüsse bezüglich ihrer semantischen Beziehungen aufgezeigt werden. Für diese Analysen müssen unterschiedliche geometrische und semantische Auflösungen der Daten berücksichtigt werden. In diesem Beitrag werden Lösungsansätze für diese Probleme vorgestellt.

1 Introduction and Overview

The growing availability of data also via the internet allows a growing interoperability of geodata as well as information sharing and reuse. This, however, presumes that the content of the data is known in order to draw meaningful and correct conclusions. Thus, for a beneficial data integration, which is a prerequisite to interoperability of data and services in Spatial Data Infrastructures (SDI), both the semantic and the geometric correspondences of these data sets have to be known.

If the semantic relationships between different geo-ontologies, like equivalence-, disjunction- or inclusion-relations are known, a geometric integration can be accomplished, e. g. a fusion or alignment process of geometry in order to obtain one improved geometry. Also an attribute transfer between the data sets is possible to enrich the

1432-8364/07/2007/0447 $ 2.75

existing knowledge about the objects. Although the semantics of the attributes still will remain unclear, the availability of this additional attributes may be helpful for new applications. The general case is that the semantic relationships between arbitrary data sources are unknown and the corresponding semantic object classes have to be determined.

In this paper we are discussing two cases. In case A, we use two data sets of similar scale where the semantic correspondences are unknown. Our approach is to use an instance-based determination of transformation rules between the ontologies. In case B we use two data sets that describe settlement objects stemming from different scales. Although the semantic relationships between the ontologies are known, a direct geometric integration is not possible due to the different granularity of the objects involved. In this case, an aggregation of the detailed data set has to be performed first in order to derive a comparable geometric object description.

The paper is organized as follows. In the next section the background of the research is sketched and references to existing work are given. Then, our methods for the two integration cases are presented, first the semantic integration of data of similar scale, then the geometric integration of data of different scales. A summary and an outlook conclude the paper.

2 Related Work

Interoperability and especially data integration faces different types of problems (Bishr 1997): it has to take structural, semantic and geometric differences in the data sets into account. Structural interoperability can be achieved using standardized data formats (e. g. ISO, OGC). The most difficult problem is semantic interoperability as it deals with the task of determining correspondences between object descriptions stemming from different user communities. In this way the corresponding semantics of the object category "lake" and "See" in an English and German topographic data set automatically have to be determined (which, of course, is not only a question of language). The general approach in the identification of semantic correspondences is to do a manual schema integration using expert knowledge together with given object catalogues or ontologies (Kokla 2006). Such a process is not adequate and not longer feasible if arbitrary data sets, e. g. downloaded from the internet, have to be integrated. Therefore, automatic techniques are needed. Rodriguez & Egenhofer (2003) use equality and similarity measures to determine relations between classes from different ontologies. Another method to automate the integration process, is a so called instance-based or extensional determination of schema transformation rules (Volz 2005, Duckham & Worboys 2005). The underlying idea of this approach is that, if two objects have an identical name and / or geometrically coincide, then they probably also have something in common on the semantic level. Duckham & Worboys (2005) use the lattice theory to determine possible class correspondences. This formally very elegant way does not take the relative frequencies of correspondences into account. This is done by Volz (2005), however, only with a manual evaluation. In our approach we want to link both approaches in order to be able to determine semantic relationsships with a corresponding probability and thus quality values.

By contrast Fonseca et al. (2006) presented a framework for measuring the degree of interoperability between geo-ontologies, which only compares the descriptions of the ontologies and not the data themselves. The drawback of this approach is, that in the general case, it can not be assumed, that the names or descriptions of objects in different data sets are the same - except objects with a unique given name like names of cities or roads. Thus using the geometric relations to infer a semantic relation seems to be more promising and open for automation.

Samal et al. (2004) take both the semantic, geometric and contextual similarity into account to derive corresponding objects in different data sets. There are also data matching approaches for data sets with different granularity (Mustière 2006, Uitermark 2001). Dunkars (2003) matches data of different scales to automatically determine the links to build a multiple representation database. Therefore a number of different measures based on topology, geometry, semantics and inter-object relationships are used to compare similarity. Bel Hadj Ali (2001) also uses different measures for assessing shape and positional quality of polygons between two different data sets.

3 Data Sets, Tasks and Approaches

For our investigations we used two different test cases with two geo-data sets each, in vector format. For case *A* (described in detail in Section 4) two data sets describing topographic objects were used: GDF (data set *A*1) and ATKIS (data set *A*2) (see Fig. 1). Whereas the GDF data (Geographic Data Files) was specially developed for vehicle navigation purposes, the German ATKIS (Authoritative topographic cartographic information system) provides topographic base data. Each object is described geometrically and semantically using the object classes and attributes described in the respective object catalogues. Both data sets are of the same scale (approx. 1 : 25.000). In Tab. 1 the hierarchical organisation of the corresponding ontologies is given in a simplified form. These presentations are not complete, only the object classes used for the analysis are listed. The two data sets are modelled differently; whereas GDF uses a two level hierarchy, followed by a further distinction using special attributes, ATKIS distinguishes three hierarchical levels, where the third level is also distinguished with attributes. Looking at the data and the structure, it is clearly visible that ATKIS exhibits a higher granularity both with respect to the object classes and the number of individual object instances. Our task is to automatically determine the semantic correspondences between object classes of the two data sets. The identification of corresponding object classes is achieved by a geometric overlay and a comparison of the frequency of the occuring object relations.

For case *B* (described in detail in Section 5) we are using two data sets with different scales and contents from China (see Fig. 2). Data set *B*1 is of scale 1 : 50.000 and specially developed for GIS purposes with comprehensive attributes while data set *B*2 is developed mainly for cartographic visualization in scale of 1 : 10.000 and contains no attributes. We are concentrating on settlement structures, namely residential areas in the small scale data and buildings in the large scale data set. The task is to transfer the attributes of the small scale data set *B*1 to the individual buildings in *B*2. Whereas on a first glance the solution to this problem might look simple, as the semantic equivalence between the object classes is already known, the situation is more complicated, because the data sets have been acquired at different time instances: data set *B*2 is more current and contains a lot of new buildings, which are not reflected in the residential areas of data set *B*1. Therefore, a mere geometric overlay or containment check of buildings and residential areas is not suffi-

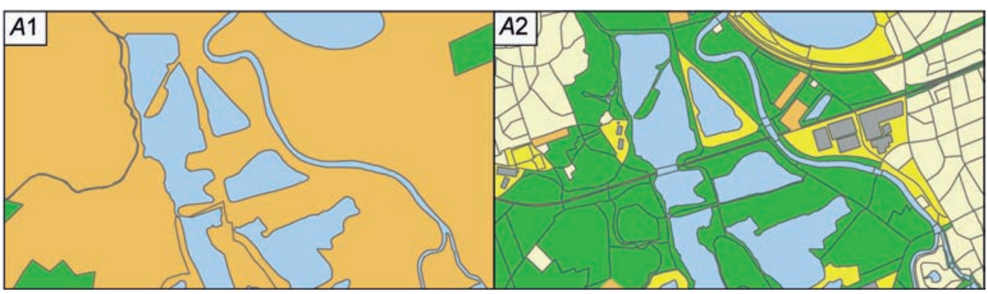

Fig. 1: Data sets for test case *A*: GDF – (data set *A*1) (*left*) and ATKIS – (data set *A*2) (*right*).

Tab. 1: The hierarchical organisation with semantically describing class names of the *A*1- (*top*) and A2-geo-ontology (*bottom*) – from superclasses (*Level 1*) to subclasses (*Level 2 or 3*) with available attributes (*right*). Σ is the total number of instances in the test area.

	Level 1	Level2	Level3		Attributes	Σ
*A*1-GDF	11 Administrative Areas	1119 Order 8				2
		. . .				
	43 Waterways	4310 Water Element			Dispclass 1, 2, 3, 4, 5	10
		. . .				
	71 Land cover and use	7120 Woodland				3
		7170 Park, Garden				1
		9715 Industrial Area				6
		. . .				
	. . .					
*A*2-ATKIS	2000 Residential Areas	2100 Built-on-areas	2101, 2111, 2112, 2113, 2114, 2126, 2132, 2134, . . .	e. g. Residential A., Industrial A.		754
		. . .				
	4000 Vegetation	4100 Vegetation Areas	4101, 4102, 4103, 4107, 4108, 4111, 4199, . . .	e. g. Forest, Moor, Farmland	FKT 2730, 9999 VEG 1000	226
		. . .				
	5000 Water Areas	5100 Water Areas	5101, 5102, 5112, . . .	e. g. Lake, River	BRG 9997, HYD 1000, OFL 11100, . . .	66
		. . .				
	7000 Regions	7100 Administrative Areas	7101, . . .	e. g. Administrative Unit		2
		. . .				

Fig. 2: Data sets for test case *B*: 1 : 50.000 – (*B*1) (*left*) and 1 : 10.000 – (*B*2) (*right*).

cient. Thus, we propose to aggregate neighboring buildings first to derive a comparable geometric representation and calculate the geometric relations to the residential areas subsequently.

4 Derivation of Semantic Relations of Data with Similar Granularity

The set up of semantic correspondences of two different data sets can be done by means of expert knowledge based on the known meaning of the terminology used by the organizations which model and capture the data sets. For example the expert could establish a semantic correspondence between $A2$-5100 (water areas) and $A1$-4310 (water element) by looking at the descriptions and definitions in the object catalogues. This inference, however, is at present only possible by a human analysis and detailed knowledge about the data. In the general case we cannot assume that expert knowledge is always available and human interaction is feasible. Therefore, the knowledge about semantic correspondences between object classes in different data sets has to be detected. Our method is to identify corresponding semantic object groups automatically by the analysis of geometric characteristics of the instances.

4.1 *Method of Geometric Overlay*

For identifying corresponding instances a geometric overlay of the data sets representing the same geometric extent is done. In the analysis we assume that the data sets are organized in layers each representing an object class. In our data sets the instances inside a layer are modelled in a tessellation, but instances of different layers within a data set can overlay. A typical example are administrative objects that often encompass larger areas and generally overlay all the other objects, so called "container objects". Accordingly a simple intersection of the data sets without consideration of further characteristics returns more than one matching candidate to an object, which could cause difficulties for the further analysis. But not only the layer structure, also the geometric discrepancies at the object boundaries themselves, occuring due to capturing by different organisations causes an increasing number of possible matching candidates, because adjacent objects may partially overlap.

For identifying corresponding objects all layers of data set $A1$ have to be intersected with all layers of data set $A2$ and vice versa. The result is a list of intersecting objects including also the relations to "container objects". That makes the procedure for large data sets with very different levels of detail or different amount of objects time consuming and inefficient. Therefore, in a preprocessing step all layers that do not overlay at all will be taken out of the analysis to reduce the data quantity for the actual overlay procedure. To this end, the instances on each layer are aggregated to one object, overlaid with each layer in the other data set and then the ratio $\mathsf{R_c}$ (similar BEL HADJ ALI 2001) between the intersection area and the union area is determined for every possible object class combination with

$$\mathsf{R_c} = \frac{|A1_i \cap A2_j|}{|A1_i \cup A2_j|} \tag{1}$$

in which $A1$ and $A2$ describe the different data sets and $\mathrm{i, j} = 1,2, \ldots$ are the occurring object classes on the lowest level of these data sets. The value of the ratio $\mathsf{R_c}$ ranges from 0 for disjunction to 1 for absolute equality. Using this ratio $\mathsf{R_c}$ a prediction about probable and improbable matching partners is not absolutely possible. Only matching candidates with a ratio $\mathsf{R_c} = 0$ are allowed to be dismissed, because even a low ratio does not necessarily mean that the matching is absolutely improbable. Due to different modelling, completeness and / or up-to-dateness of the data sets very different values may occur. By excluding all combinations, which fell below the value $\mathsf{R_c} \leq 0{,}01$ the results of case A show a reduction from 95 (5×19 matrix) object class combinations to 32. Among those, four complete $A2$-classes had no corresponding $A1$-classes and are no longer needed to be processed.

To further reduce the number of matching candidates an exclusion of neighbouring or minimal overlapping objects is done by introducing and considering the geometric criterion area, especially the object size and intersection area, in the analysis. In this process the overlay ratios of the size of the objects O and the intersection area I are calculated as follows

$$\mathsf{R}_{\mathsf{O},A1} = \frac{I \cdot 100\,\%}{\mathrm{O}_{A1}} \quad \text{and}$$
$$\mathsf{R}_{\mathsf{O},A2} = \frac{I \cdot 100\,\%}{\mathrm{O}_{A2}}. \tag{2}$$

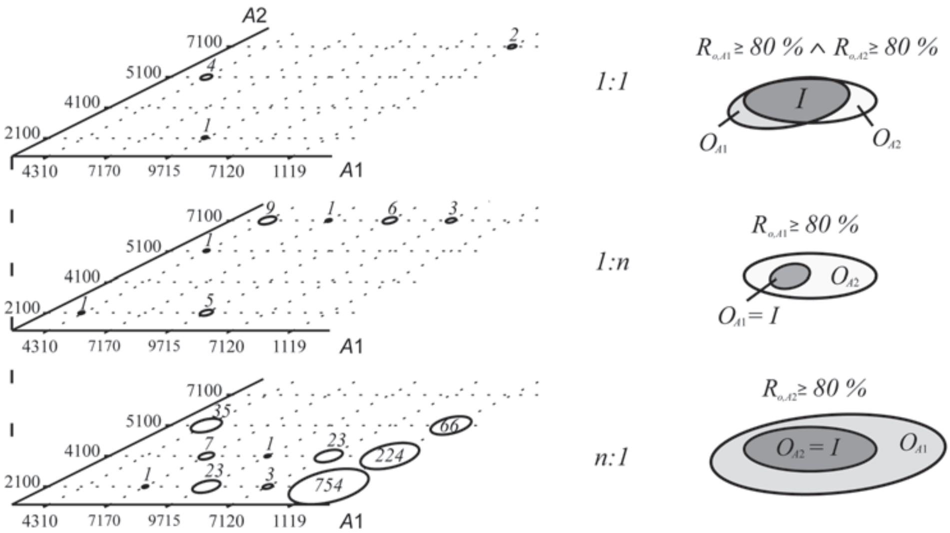

Fig. 3: Results of the overlay method displayed by a frequency matrix: level 1 – contains the matching candidates which fulfill the condition of 1 : 1 relation, level 2 and level 3 contain the candidates which fulfill the condition of 1 : n relation. The figures at the circles indicate the number of possible matching candidates.

Taking small geometric differences into account, we consider objects to match, when the ratio is 80 % or better. In the case of a 1 : 1-relationship following condition has to hold: $R_{o,A1} \geq 80\,\% \wedge R_{o,A2} \geq 80\,\%$.

But if the data sets are differently modelled, i. e. one data set contains more aggregated objects than the other data set, the search for 1 : 1 relations returns only few objects. Therefore the analysis is extended from 1 : 1 relations to 1 : n relations. In that case the condition $R_{o,A1} \geq 80\,\% \vee R_{o,A2} \geq 80\,\%$ has to be fulfilled. As a result a three-level frequency matrix can be prepared that displays the amount of the remaining matching candidates as illustrated in Fig. 3 (left). In order to ensure the readability the class levels 3 of the data set *A*2 are summarized to the respective level 2. On the right side of Fig. 3 the conditions are clarified with the schematic diagrams, that lead to the results presented on the left side.

In the next section the analysis of semantic correspondences from these individual results of the single object classes on the lowest levels is described, which is done up to now in a manual examination.

4.2 *Analysis of Relations for Semantic Integration*

The results of the geometric overlay method on instance level presented in section 4.1 have to reveal the correspondences on class level.

Firstly, clear relations can be derived. For example 1 : 0-disjunction relations ($A1_i \cap A2_j = \emptyset$) are detected between object classes *A*1-4310 and *A*2-4100 respectively *A*1-7120, *A*1-7170, *A*1-9715 and *A*2-5100. A 1 : 1-equivalence relation ($A1_i \equiv A2_j$) between the classes *A*2-7100 and *A*1-1119 is detected, because all (in this case two) objects of both object classes meet the condition $R_{o,A1} \geq 80\,\% \wedge R_{o,A2} \geq 80\,\%$. Additionally to the clear relations also a great amount of matching candidates is found in level 2 and 3 of Fig. 3 representing examples for the 1 : n-inclusion condition ($A1_i \subseteq A2_j$). But not all of these candidates represent true 1 : n relations, because all relations to "container objects" are also included. These large objects (administrative areas) contain nearly all objects of the other data set, i. e. objects of *A*2-7100 and *A*1-1119. These

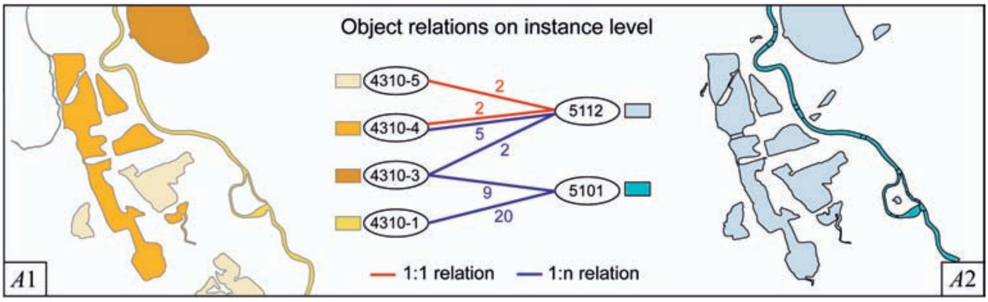

Fig. 4: Object relations on instance level of the object class *A*1–4310 with attributes (left) and *A*2–5101, –5112 (right). The red lines display 1 : 1 relations and the blue lines 1 : n relations with the number.

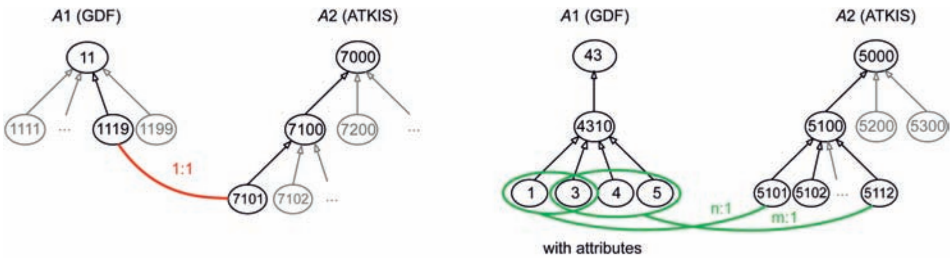

Fig. 5: Resulting semantic relations between object classes of both data sets derived from the instances-relations.

matching candidates have to be dismissed from the further analysis, because they are no real matching partners between corresponding classes. The remaining matching candidates are analysed in more detail. For example between object classes of *A*2-5100 and *A*1-4310 all kinds of instance relations exist, which is an indicator for semantic correspondences. If the hierarchy of the ontologies is regarded on the lowest class level a 1 : n relationships between object classes *A*1-4310 and *A*2-5101, *A*2-5112 exist. In order to be able to identify more detailed relations, in the further analysis additional attributes, i. e. attribute values like proper names or geometric properties will be included. This may lead to better distinctions, because in many cases the relevant semantic information is expressed by these attributes.

Including the available attributes of the object classes *A*1-4310 in the analysis the following relations between the instances can be detected: objects of object class *A*1-4310 with attribute values 1 (*A*1-4310-1) and 3 (*A*1-4310-3) match to objects of *A*2-5101 and *A*1-4310-3, -4 and -5 match to object class *A*2-5112. Whereas the former has only 1 : n relations the latter contains also 1 : 1 relations that have to be more important to the evaluation than 1 : n relations. Object class *A*2-5102 has no matching partners at all in this test area. The object class *A*1-4310-3 matches to *A*2-5101 and also to *A*2-5112 (see Fig. 4).

These derived semantic relations between the object classes of both data sets can be represented with the help of the ontology hierarchies. Therefore in Fig. 5 the results for the 1 : 1 equivalence relation between the lowest levels of *A*1-11 and *A*1-7000 and the relations between level 2 of *A*1-43 with attributes and level 3 of *A*2-5000 are presented; from the latter, also a higher level correspondence between *A*1-4310 and *A*2-5100 can be derived.

Beside the existing attributes also derived attribute values of the geometry can improve the analysis of the semantic correspondences in the future. As geometric properties would be conceivable compactness, elongation, rectangularity, which are typical for certain objects. For example buildings in general have parallel borders and right angles whereas lakes or ponds have in most cases a round, organic outline. Furthermore the integration of context-dependent relationships will be considered. For example an object X of *A*1 corresponds to an object Y of *A*2 only when X is in context C1 and Y in context C2. In another context, X would correspond to another class.

5 Attribute Transfer of Data with Dissimilar Granularity and Up-to-dateness

5.1 *Aggregation and Matching*

For identifying corresponding objects in data with dissimilar granularity a simple geometric overlay of the data sets does not always lead to the desired result. Objects in large scale data sets have more geometrical details than objects in the small scale. Objects (in our case buildings) in border areas in the large scale data set may clearly belong to the group of buildings in their local environment although they might not fall within the corresponding settlement area in the smaller scale. This is even more true, as in our case due to temporal differences, new buildings are present in the detailed data set, that are not reflected by residential areas in the smaller scale data. In this test case the knowledge about the semantic corresponding object classes is presumed to be known, namely the aggregation relationship of buildings from *B2* and settlement areas in *B1*. As described above, the goal of the analysis is the identification of the individual corresponding objects in order to accomplish a transfer of semantic information (e. g. name attributes) from *B1* to *B2*. To establish the relationships between these objects a three-step-approach is proposed. With no additional information available (e. g. road network) only geometric reasoning can be used for the analysis.

The first step involves a distance-based aggregation of objects of data set *B*2. By means of this procedure a generalization is accomplished, in order to overcome the granularity differences. For this purpose a buffer is computed around the buildings. The determination of the buffer distance is crucial. The most suitable buffer distance can be determined depending on the building density. To this end clustering approaches can be used (e. g. ANDERS 2004). The aggregation of the building objects of *B*2 is illustrated in Fig. 6 (left). In this first experiment, we used a fixed buffer size of 10 m derived from experiments.

The second step involves the search for correspondences of the buffered building polygons (bps) and the residential areas (ras) of *B*1. We used a distance-based approach by determining Voronoi cells (Vcs) of the residential polygons of data set *B*1. These polygons, as illustrated in Fig. 6 (middle) divide the whole area into seamless cells and every cell is the best range corresponding to a 1 : 50.000 ra object. In this figure the red lines are Voronoi polygon boundaries and the pink blocks are residential objects in 1 : 50.000. In the third step a geometric overlay with the new bps of 1 : 10.000 and the Vcs of 1 : 50.000 is done (see Fig. 6, right).

To analyze the correspondences of bps and Vcs two methods can be used. One method is to calculate the centroid point of the bps. If the point is within one Vc, we decide that the bp belongs to this cell polygon, and thus to the corresponding ra. In Fig. 7 (left) the centroid points of bps and Vcs are displayed. In general case the mapping is clear, but at the border of the Vcs the result is not always satisfying. The bp in the red circle will be mismatched to the Vc above.

To improve the result, another method is to decide by means of the degree of overlap. If the larger part of the bp is within a Vc, the polygon is considered to belong to this cell. That means for the example displayed in Fig. 7 (right), that the polygon must be assigned to the lower cell, because the larger part displayed in dark grey is located there.

Fig. 6: Three-step-approach for establishing relationships between objects of data sets with dissimilar granularity; Creation of buffer polygons (bps) (orange) (*left*), Voronoi cells (Vcs) (red) of 1 : 50.000 (*middle*) and Result of overlay of bps and Vcs (*right*).

Fig. 7: Mapping results of the centroid point method (*left*) and the degree of the overlapping areas (*right*). The blue crosses are the centroids of the orange buffer polygons.

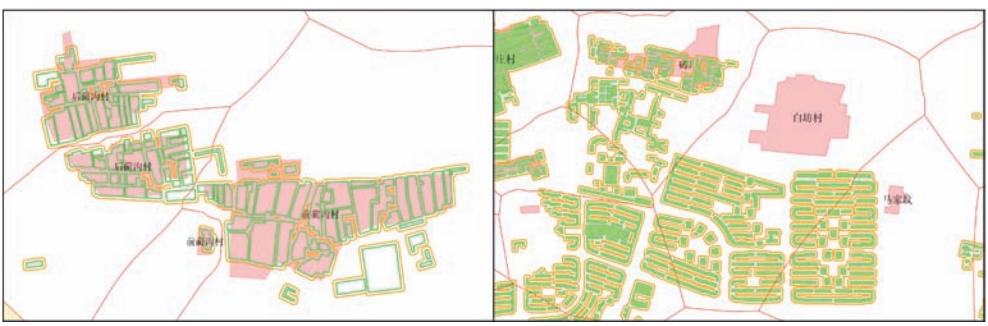

Fig. 8: Wrong assignment of corresponding residential areas due to aggregation (*left*); different spatial situation due to different acquisition time (*right*).

With this matching of bps to Vcs the buildings inside the bps are also assigned.

Furthermore, the degree of overlap between bps and Vcs and ras can give an indication to the quality of the correspondence found. The following analysis steps can be performed: first of all, the degree of overlap of bp and Vc gives an indication to the containment in a cell and also to the fact, whether the buildings have been correctly aggregated. If the degree of overlap of bp and Vc is smaller than 50 % this means that half of the bp-area is located in one or more other Vcs. This could indicate that the buffer distance had been set too large (see Fig. 8, left). The degree of overlap of bp and ra indicates, how well these two areas fit; the closer the value is to 1, the better is the overlap and thus also the reliability of the assignment of the attributes.

For rapidly changing areas, this approach can not identify corresponding buildings correctly because the spatial situation has changed dramatically: a lot of new buildings have been built, which cannot reliably be assigned to the old residential areas. However, the measure of degree of overlap between both bp and Vc and bp and ra can give an indication and help to perform a visual inspection of the quality of the automatically determined correspondences easily.

5.2 *Attribute Transfer*

The results of the matching are corresponding building objects and small scale residential objects. The residential areas of 1 : 50.000 have the attribute place name, which can now be transfered to the buildings 1 : 10.000, in order to enrich their mere geometry. The transfer of the attributes can be achieved by joining their attribute tables. The quality of the overlap can be taken into account, in order to qualify the reliability of the transferred attributes.

6 Discussion of Results and Future Work

The presentation describes ongoing work on semantic data interoperability. The need of integrating data sets of different origin and different granularity is evident, especially for a data reuse. The work on semantic integration is in a very early stage. In the future, there will be a focus mainly on the improvement of the identification of semantic correspondences by an extended analysis, mainly by automating this process. To this end, additional attributes to the comparison of the data, like shape parameters, will be introduced. Furthermore the analysis of correspondences between the object classes on lowest level must be extended, because equivalent concepts may be expressed at different levels in the hierarchy of the ontology. The analysis will be automated using association rules from Data Mining (Agrawal & Srikant 1994).

The matching of objects of two data sets with dissimilar granularity is more successful, if they share a similar up-to-dateness. In the analysed example of an urban or suburban area in China the changing of the areas was so rapid and far reaching, that the proposed three-step-approach was not suitable for some parts of the test area; it seems to be more appropriate for rural areas, where changes do not happen so often. To improve the results in urban and suburban areas the introduction of additional information about possibly partitioning objects, e. g. road and river networks, is essential. Another issue is a refined aggregation process, which should take the structural similarity of the buildings (large industrial vs. small residential buildings) into account.

Acknowledgements

This research is funded by the German Science Foundation (DFG) and National Nature Science Foundation of China (NSFC): 40620130438. The support is gratefully acknowledged.

References

Agrawal, R. & Srikant, R., 1994: Fast algorithms for mining association rules. – Proceedings of the 1994 International Conference for Very Large Data Bases (VLDB'94), 487–499.

Anders, K.-H., 2004: Parameterfreies hierarchisches Graph-Clustering-Verfahren zur Interpretation raumbezogener Daten. – Dissertation, Fakultät Luft- und Raumfahrttechnik und Geodäsie der Universität Stuttgart, Stuttgart.

Bel Hadj Ali, A., 2001: Positional and shape quality of areal entities in geographic databases: quality information aggregation versus measures classification. – Proceeding of ECSQARU'2001 Workshop on Spatio-Temporal Reasoning and Geographic Information Systems.

Bishr, Y., 1997: Semantic aspects of interoperable GIS. – Wageningen Agricultural University and International Institute for Aerospace Survey and and Earth Science (ITC), Enschede.

Duckham, M. & Worboys, M., 2005: An algebraic approach to automated geospatial information fusion. – International Journal of Geographical Science **19** (5): 537–557.

Dunkars, M., 2003: Matching of Datasets. – Proceedings of the 9th Scandinavian Research Conference on Geographical Information Science, p. 67–78.

Fonseca, F., Câmara, G. & Monteiro, A.M., 2006: A Framework for Measuring the Interoperability of Geo-Ontologies. – Spatial Cognition and Computation **6** (4): 309–331.

Kokla, M., 2006: Guidelines on Geographic Ontology Integration. – ISPRS Technical Commission II Symposium, p. 67–72.

Mustière, S., 2006: Results of experiments on automated matching of networks at different scales. – Proceedings of the ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data, 92–100.

Rodriguez, A. & Egenhofer, M., 2003: Determining Semantic Similarities Among entity classes from different ontologies. – IEEE Transactions on Knowledge and Data Engineering, **12** (2): 442–456.

Samal, A., Seth, S. & Cueto, K., 2004: A feature-based approach to conflation of geospatial sources. – International Journal of Geographic Information Science **18** (5): 459–490.

Uitermark, H.T., 2001: Ontology-based geographic data set integration. – PhD Thesis, Computer Science Department, University of Twente, Enschede, The Netherlands.

Volz, S., 2005: Data-Driven Matching of Geospatial Schemas. – Cohn, A.G., Mark, D.M. (eds.): Spatial Information Theory. Proceedings of the International Conference on Spatial Information Theory (COSIT '05), Lecture Notes in Computer Science 3693, Springer, 115–132.

Addresses of the Authors:

Dipl.-Ing. Birgit Kieler, Prof. Dr.-Ing. habil. Monika Sester, Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik, Appelstraße 9A, D-30167 Hannover, Tel.: +49-511-762-3588, Fax: +49-511-762-2780, e-mail: {Birgit.Kieler, Monika.Sester}@ikg.uni-hannover.de

Msc. HaiQing Wang, Dr. Jie Jiang, National Geomatics Center of China, 1 Baishengcun, Zizhuyuan, Beijing, 100044, Tel.: +89-10-68483218, Fax: +86-10-68424101, e-mail: {whq, jjie}@nsdi.gov.cn

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007