METHOD FOR DERIVING SPATIAL DATA OF DRESDEN TO SMALLER SCALES

Karsten Pippig and Dirk Burghardt

Karsten.Pippig@tu-dresden.de, Dirk.Burghardt@tu-dresden.de

Technische Universität Dresden, Institute of Cartography, Helmholtzstraße 10,

D-01069 Dresden, Germany

Abstract
Cartographic generalisation is one of the most pivotal issues in cartography. From the 1960s on, a development has taken place from free practical map generalisation depending on the abilities of the mapmaker towards a computer assisted automation. This paper summarizes the conception and implementation for an application of the automatic generalisation for data of the General Map provided by Dresden's municipal survey office and for the ATKIS® base DLM outside the municipal area. The focus is set on the selection and implementation of appropriate generalisation methods to form an overall process as well as the implementation of an application for model generalisation. The result is a stand-alone application being capable of supporting Dresden's municipal survey office in generalising its data.

Keywords: cartography, model generalisation, generalisation operators

1 Introduction

The municipal survey office of Dresden supplies spatial data in several analogue and digital forms mapping the entire city. The Digital General Map is one of the main products. It represents the classified traffic network (ESKN) with official street names and built-up areas, waterways, parks, and selected public buildings as well as the official district boundaries and names (DRESDEN, 2011). The data covering the surrounding communities are provided by the ATKIS® base DLM of the land surveying office of Saxony (GeoSN). The Digital General Map forms the basis for many GIS applications and derived thematic maps of Dresden authorities. In 2010, for example, Dresden published a municipal atlas in 1:90,000[1] but using the corresponding spatial data of the Digital General Map in 1:25,000. Indubitably, it leads to presentation and legibility problems in representing these data. Therefore, it is necessary to derive the data for smaller scales (1:50,000 and 1:90,000) and to represent them generalised for cartographic purposes.

2 Data models of Dresden's municipal survey office

2.1 ESKN (Erweitertes Straßenknotennetz; eng: Extended Road Junction Network)

ESKN is the database containing the entire traffic network. It is used to query for specific streets and to represent the road, railroad, and tram network as the map base for all small-scale thematic maps. It comes in two reference layers: 1:5,000 and 1:25,000; whereas the traffic network is generalised in the 25k-layer (incl. collapsed and displaced road segments).

2.2 EBK (Erweiterte Blockkarte; eng: Extended City Block Map)

EBK is the database that represents the urban subdivision of Dresden based on city blocks. The city blocks are derived from the traffic network (ESKN), district boundaries and the river Elbe with particular consideration of the specific land use. For statistical planning and cartographic purposes, each block is classified by statistical and administrative criteria with three different detailed types: coverage type (e.g. open coverage type), specific land-use type (e.g. residential area), and building type (e.g. terraced house).

3 Methodical approach

The entire process of automatic generalisation is designed as a condition-action model according to the characterisation of HARRIE & WEIBEL (2007). The various generalisation operators are carried out sequently after each other (Figure 1). Each generalisation step is preceded by structure recognition (compare to BRASSEL & WEIBEL, 1988). It may refer to the semantics, topology, or geometry of an object. Within each generalisation process, rules for the particular object structure are defined, describing whether and how an object has to be generalised. The parameters for the conduct are

[1] The scale of 1:90,000 is chosen to represent Dresden in its full expansion on an A3 sheet.

stored in a database and retrieved during the process. The choice of the generalisation operators and the optimal sequence, as it is shown in figure 1, result from several empirical tests during the implementation phase.

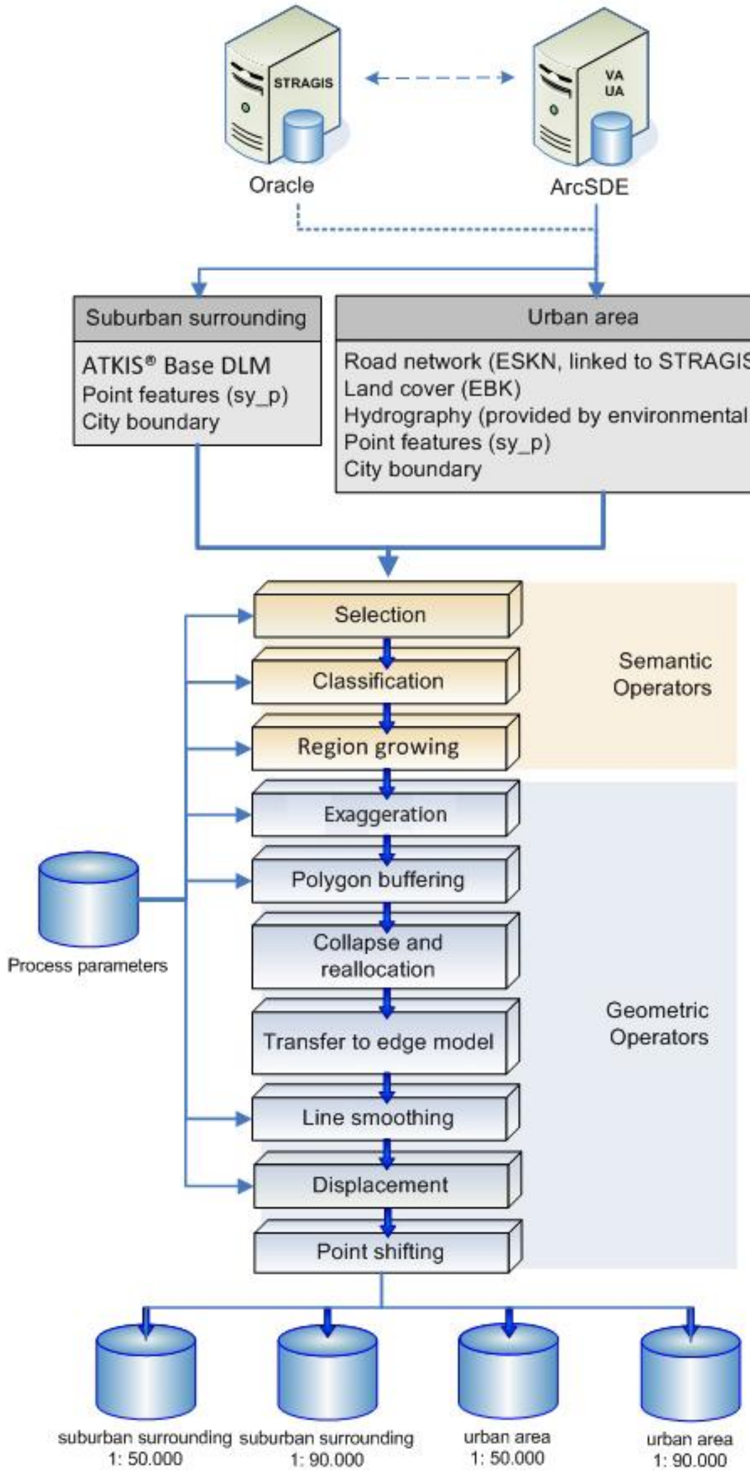

Figure 1. Overall flow of the generalisation process

Due to the data sources' different structures, it is reasonable to execute the generalisation for the urban data of Dresden and for surrounding areas separately.

At the beginning, semantic operators are executed. These include selection, classification, and fusion. During the execution of the selection and classification, different structures of the various data sources have to be considered. The data are transferred into a single unified structure and into equal object classes, allowing for the following operators to use this common structure. The execution of semantic operators at the beginning aims at reducing the density. For example, it is possible to reduce the number of polygons within the urban area from 15,769 to 3,418 only by using classification and fusion algorithms. The geometric operators can be divided into exaggeration, collapse, line smoothing, and displacement. Here, the geometric operations may only be performed when considering topological relationships to other object classes. This work focuses on the programmatic implementation of collapsing small areas followed by reallocation of adjacent polygons.

4 Generalisation operators

4.1 Selection and classification

The selection and the classification of polygon features are designed in a matter to represent the land use over the entire urban or surrounding area completely and comprehensively whereby the semantic feature types are merged to higher classes. The road network is pruned and only represents primary and secondary roads. The selection and classification of ATKIS® data in the surrounding area follow the data's structure used in the urban area of Dresden.

4.2 Region growing

The fusion of objects in terms of region growing, as it is described by HAUNERT (2008), is carried out for polygon features being smaller than the minimum area. The latter is defined as 1 ha in 1:50,000 and 3.24 ha in 1:90,000. These detected features are fused with an adjacent polygon feature. The assignment of a relevant to an adjacent feature is done semantically by applying a priority table (appendix, Table 1). This table defines the relationship of two object classes in each case. As to the content, the priority table follows similar tables used for example in the CORINE Land Cover 2000 project (KEIL, KIEFT & STRUNZ, 2005) or by YOALIN, MOLENAAR & KRAAK (2002). Moreover, individual priority tables by different participants were set up and their results were taken into account. However, a purely semantic fusion linked to class relationship does not lead to the desired success. Therefore, different priorities are combined according to the class relationship and length of the adjacent features' common boundary. Multiplying the boundary's relative length and the class relationship provides a quantitative value for comparing all adjacent features.

Adjacent areas that are separated by a road feature are excluded from the fusion. The same applies to areas being enlarged in the next generalisation step.

4.3 Exaggeration

Too small but significant polygon features cannot be omitted by applying generalisation. Therefore, it is necessary to enlarge them to the minimum dimensions. As an automatic assessment of each polygon feature's importance is difficult to implement, the user has to evaluate the signification individually and has to add each area of the EBK that need to be exaggerated in a separate table by its specific ID number. The medial axis of each recorded polygon feature can be determined and buffered according to a predetermined minimum width. The buffered medial axis is fused with the initial geometry of the feature. An exaggeration of individual features, on the one hand, reduces the size of adjacent features on the other hand. Overlapping areas between enlarged and adjacent polygons will be deducted from the adjacent area.

4.4 Collapse and reallocation

A collapse of areas leads to the disappearance of entire features and to a reallocation of the disappearing areas to adjacent features (HAUNERT & SESTER, 2008). This causes strong changes in the shape of the objects. It is therefore desirable to change only the geometry of features that fall short of a minimum width value. SU et al. (1997) describe morphological operators for a raster-based generalisation that can also be applied to generalise vector data. By an inward buffering (erosion) and subsequent outward buffering (dilatation), all polygon features can be split into area parts that do not meet the minimum requirements. Areas smaller than the minimum width are collapsing. The number of polygons collapsing increases with the buffer width. A minimum width of 2 mm in the map (100 m in 1:50,000 and 180 m in 1:90,000) appeared to be inappropriate (Figure 2). They led to severe and adverse changes in the shape of landscape. In practice, widths of 40 m and 70 m turned out to be more appropriate. The division of the initial polygons can cause polygon parts that may satisfy the specified minimum width, but do not meet the requirement of a minimum area of 1 ha and 3.24 ha. A collapse has to be calculated for these parts as well.

After the medial axis has been determined, all adjacent areas have to be projected onto the axis starting from the two edge nodes of the common boundary. But in case the collapsed polygon adjoins road features, the adjacent polygons cannot be projected beyond the roads. In this case, the adjacent polygons shall not be extended to the medial axis but on the road axis. This may cause overlaps or gaps if the polygon geometry is unfavourable. Another special case can occur when two separate roads adjoin the collapsed polygon. An approximately orthogonal line has to be constructed between both road segments to form a new boundary for the adjacent polygons.

The collapse and reallocation can be implemented in three steps. In the first step, the adjacent polygons are projected on the medial axis as it is described above. Then, all features are clipped to the road network. If reference points have been generated within the initial polygons previously, all polygons without a reference point can now be identified. These polygons will be collapsed in the second step and the adjacent features are projected on the road features. Analogously, the third step is similar, but the adjacent polygons will be projected on the municipal boundary, which is also the boundary between the different data.

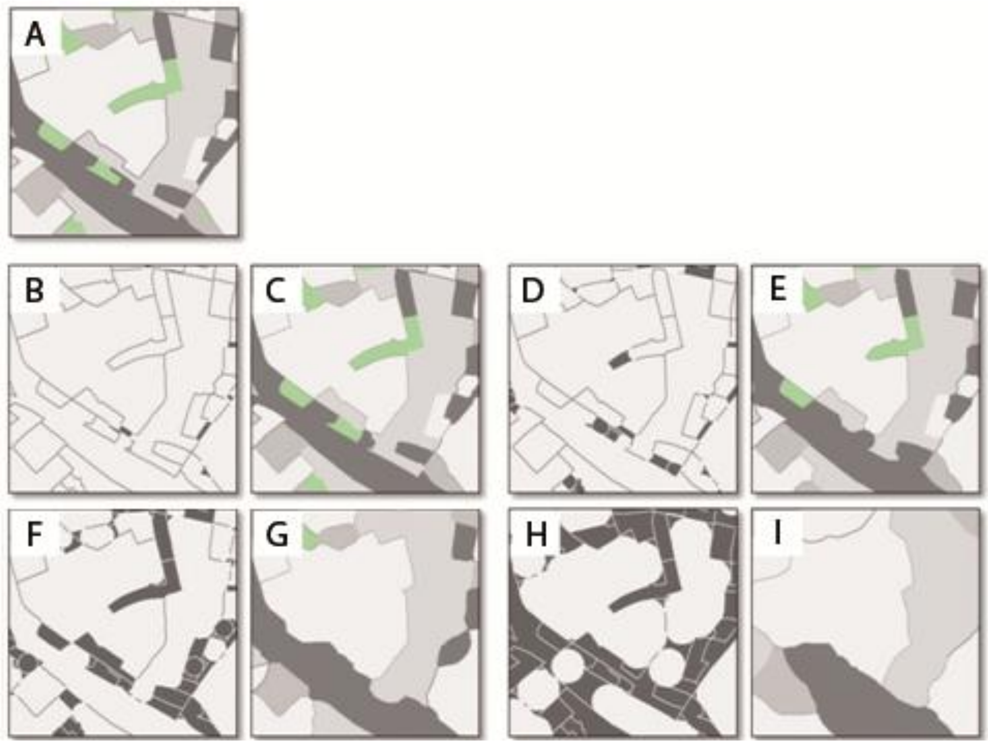

Figure 2. Results of the collapse and reallocation with different buffer sizes.

A - Initial situation, B/C - buffered with 20 m, D/E - buffered with 35 m, F/G - buffered with 50 m, H/I - buffered with 90 m. Dark areas in B, D, F and H show areas that do not meet the minimum width

4.5 Line smoothing

Line smoothing and simplification are key components of all generalisation processes and connected to the aim of adapting the line structures to scale without changing the character of the landscape. A separate smoothing or simplification of polygon and line features would not reach the desired aim, since each polygon has different start and end nodes. To avoid this, all polygon features are transformed into a model based on its topological edge-node structure. In addition, all line features (road data, municipal boundary) are transmitted to the same edge-node model. Simplification or smoothing algorithms can be applied to the edges of the model. The municipal boundary remains unsmoothed, as it is the administrative referenced boundary line between the urban and the surrounding area.

Several line smoothing and simplification algorithms have been proved in several tests. McMaster's Slide Averaging Algorithm (McMASTER, 1989) is proved to be a suitable smoothing algorithm that meets the demands of the generalisation of polygon features with regard to road data. Three neighbouring points as the number of neighbour parameter and a displacement value of 50% were taken into account. The river network that is not considered in the edge-node model; it is simplified separately. From own empirical studies of different algorithms, it can be concluded that a combination of the Douglas algorithm with a tolerance of 20 m and a subsequent smoothing by the McMaster algorithm produces optimal results.

4.6 Displacement

A suitable displacement approach is currently not integrated in the application – even though there is a need for its implementation. Attempts to carry out a displacement based on an adaptation of Nickerson's algorithm in FME were dissatisfied. A displacement approach based on energy minimization by snakes (compare BURGHARDT & MEIER, 1997) is desired, but does not provide node shifting, which is disadvantageous to retain the shape. One solution could be a displacement algorithm based on elastic beams (compare BADER & BARRAULT, 2001). A displacement approach would increase the distance between the edges of the edge-node model to the minimum distance.

4.7 Point shifting

Having applied the line smoothing and displacement operators, all edges and nodes are transferred to their original feature classes. Following this approach, it is possible to protect the topological relationship of all features. However, it is necessary to orient the point features towards the generalised line features. Currently, the point features will be reoriented on the smoothed edge by the angle and the relative distance with respect to the original edge.

5 Technical implementation

The implementation of the generalisation application is done in Visual Basic with the inclusion of ArcObjects and FMEObjects. ArcGIS provides a variety of tools and functionalities in the field of map production, spatial analysis, data management, and geoprocessing. These functionalities are available in libraries of ArcObjects for further development. Nevertheless, there are some deficits in various generalisation operations, in particular concerning the displacement and area collapse. ArcGIS 10.0 provides tools to propagate displacement and to collapse dual lines to centrelines; however, they are not usable for this application. FME, instead, offers solutions in this field. It is therefore obvious to use the functions implemented in FME for the application as well and to take advantage of both GIS platforms.

6 Evaluation

The derivation of the data in 1:50,000 produced 122 incorrect generalised features out of a sample of 3,194 polygon objects and thus necessitated a manual post processing. These features are classified with 'indefinite' during the entire process and cover 3.8% of all polygon features in the map. Thereby, the incorrect features account for a total area of 1.38 ha only (approximately 0.04% of the total area of Dresden).

After the entire process, the amount of polygon features was reduced from 6,525 to 3,071. This corresponds to a ratio of 0.471 and achieves approximately the theoretical value of 0.5 of the radical law by TÖPFER & PILLEWIZER (1966). The average polygon area has been increased by the factor 1.92. The reciprocal value of 0.52 approximately corresponds to the ideal value of the radical law, too. The increase or loss of area for all feature type in relation to the total area of Dresden is below one percent. The general size and number of polygon features are adjusted in respect to the scale without affecting the relative distribution.

In regard to the compliance with the minimum size of 1 ha (+/- 5% tolerance), it can be stated that 2,904 of all 3,071 polygon features meet this requirement; 167 fail.

Only 1,655 features meet the requirement of a minimum width of 100 m. 1,416 features undermine this requirement. However, it should be noted that a displacement approach is not currently implemented.

7 Conclusion and outlook

The resulting generalisation application provides initial approaches for the automatic derivation of data from Dresden's municipal survey office. The obtained results are rated positively by Dresden's municipal survey office and are used in map production, although a further development and improvement of the application and the algorithms would be desirable. A displacement algorithm is currently not implemented either. The lack of an adequate displacement algorithm reduces the quality of the entire generalisation.

Furthermore, the reorientation of point features has to be improved. Currently, the point features are oriented and aligned to the nearest line feature. As a consequence, a point feature is strongly offset if its distance to a line feature increases. An orientation to all surrounding line features seems reasonable and could be carried out via a triangulation. In parallel, point features have to be a displaced to avoid their overlapping.

The previously mentioned problem regarding the exaggeration is less satisfactorily solved. In particular, a suitable solution has to be sought to prevent any overlapping if two adjacent features have to be exaggregated. Presumably, this problem can be solved iteratively.

A further development of this application should be made in regard to an automated label placement and automated map generating.

References

Bader, M. & Barrault, M. (2001). Cartographic Displacement in generalisation: Introducing Elastic Beams, 4th workshop on progress in automated map generalisation, August, Beijing, China.

Brassel, K. E. & Weibel, R. (1988). A Review and Conceptual Framework of Automated Map generalisation, *International Journal of Geographical Information Systems* **2**(3), pp 229-244.

Burghardt, D. & Meier, S. (1997). Cartographic Displacement Using the Snakes Concept, In *Semanitic Modeling for the Acquisitaion of Topografic Information from Images and Maps:* W. Förstner & L. Plümer (eds.), Birkhäuser Verlag, Basel, pp 59-71.

Harrie, L. & Weibel, R. (2007). Modelling the Overall Process of generalisation, In *Generalisation of Geographic Information: Cartographic Modelling and Applications:* W. Mackaness, A. Ruas & A. T. Sarjakowski (eds.), Elsevier, Oxford, pp 67-87.

Haunert, J.-H. & Sester, M. (2008). Area Collapse and Road Centerlines based on Straight Skeletons, *Geoinformatica* **12**(2), pp 169-191.

Haunert, J.-H. (2008). *Aggregation in Map generalisation by Combinatorial Optimization*. PhD Thesis, Leibniz Universität Hannover, Germany.

Keil, M., Kiefl, R. & Strunz, G. (2005). *CORINE Land Cover 2000 – Germany*. Final Report. On behalf of the German Federal Environmental Agency (UBA), German Aerospace Center (DLR), Oberpfaffenhofen, Germany.

Su, B., Li, Z., Lodwick, G. & Müller, J.-C. (1997). Algebraic models for the aggregation of area features based upon morphological operators, *International Journal of Geographical Information Science* **11**(3), pp 233-246.

McMaster, R. B. (1989). The integration of simplification and smoothing algorithms, *Cartographica* **26**(1), pp 101-121.

Töpfer, F. & Pillewizer, W. (1966). The principles of selection, *The Cartographic Journal* **3**(1), pp 10-16.

Yoalin, L., Molenaar, M. & Kraak, M.-J. (2002). Semantic similarity evaluation model in categorical database generalization, In proceedings of the joint ISPRS commission IV symposium, Spatial Data Handling and 95th annual CIG geomatics conference: Geospatial theory, processing and applications, ISPRS commission IV - WG IV/3, July, Ottawa, Canada. CD-ROM, p 7.

Links

Dresden (2011). *Übersichtskarte Dresden*, http://www.dresden.de/de/02/06/02/c_04.php [23.02.2011]

Appendix

Figure 3a. Excerpt from the initial Digital General Map of Dresden in 1:25,000 (by kind permission of Dresden's municipal survey office)

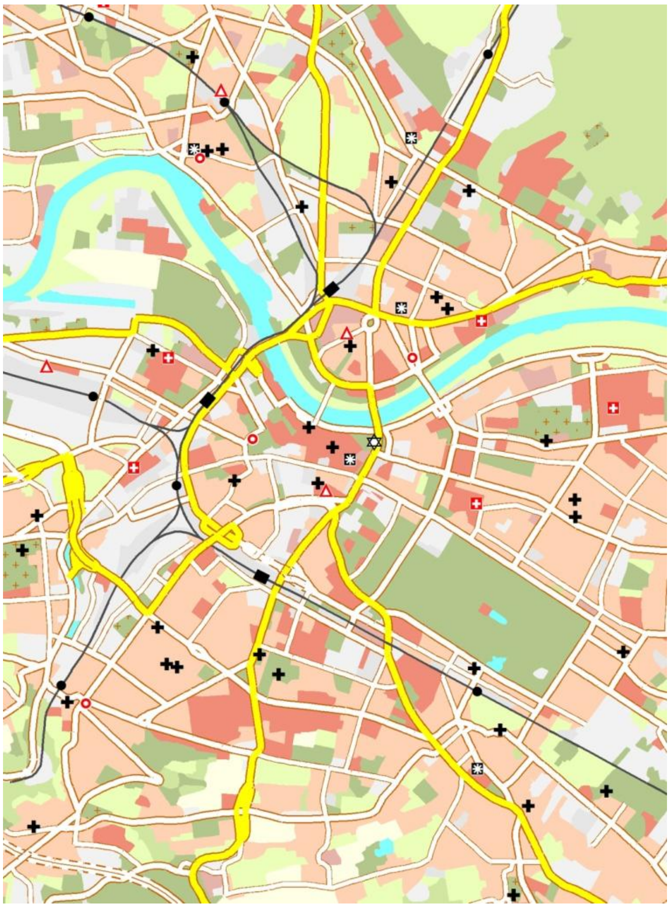

Figure 3b. Excerpt from the generalised map in 1:50,000

	Residential area	Mixed-use area	Special building area	Industrial/commercial area	Disposal site	Green area	Cemetery	Grassland	Swamp	Vineyards	Fruit cultivation/gardening area	Farmland/ fallow land	Forest/wood	Water area	Traffic area	indefinite
Residential area	100	90	90	80	70	80	80	70	50	70	70	70	60	40	80	0
Mixed-use area	90	100	90	90	70	80	80	70	50	70	70	70	60	40	80	0
Special building area	90	90	100	80	70	80	70	70	50	70	70	70	60	40	80	0
Industrial/ commercial area	70	70	70	100	90	60	50	40	40	40	40	50	40	30	90	0
Disposal site	60	60	60	90	100	50	40	50	50	40	40	50	30	30	70	0
Green area	80	70	70	60	60	100	80	50	40	50	60	50	50	40	60	0
Cemetery	70	70	60	60	50	80	100	60	40	40	40	40	50	30	50	0
Grassland	60	60	60	40	40	70	60	100	70	70	70	80	40	30	40	0
Swamp	30	30	30	30	30	30	30	60	100	40	50	60	60	70	30	0
Vineyards	60	60	60	50	40	50	40	70	40	100	90	70	50	30	40	0
Fruit cultivation/ gardening area	60	60	60	50	30	50	40	70	30	90	100	70	70	30	40	0
Farmland/fallow land	60	60	60	60	60	50	40	80	50	70	70	100	60	30	50	0
Forest/wood	40	40	40	40	40	60	50	50	50	40	60	60	100	30	40	0
Water area	40	40	40	40	40	50	40	50	70	40	40	40	40	100	40	0
Traffic area	80	80	80	90	70	60	50	50	40	50	50	50	40	30	100	0
indefinite	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0

Table 1. Priority table for the determination of class relationships