A Road Network Selection Process Based on Data Enrichment and Structure Detection

Guillaume Touya
COGIT Laboratory, Institut Géographique National, 2 av. Pasteur, 94160 St-Mandé, France
guillaume.touya@ign.fr

ABSTRACT. The presented paper deals with a generic process for road network selection based on data enrichment and structure detection. The first step is to detect significant structures and patterns of the road network like roundabouts or highway interchanges. It allows to enrich the initial dataset with implicit geographic entities and structures. Using this enrichment, the following step is the selection of roads in rural areas thanks to graph theory techniques. After that, urban roads are selected by means of a block aggregation complex algorithm. Continuity between urban and rural areas is guaranteed by a method based on the notion of strokes. Finally, some previously detected structures are typified to maintain their properties in the selected network. This automated process has been fully implemented on Clarity™ and tested on a large dataset.
KEYWORDS: generalisation, selection, road network, structure detection

1. Introduction and Related Work

In geographic information systems, generalisation is a process that aims at summarizing geographical information from a geo-database in order to produce a less detailed database or map. When the goal is map production, it is called cartographic generalisation and it consists in solving cartographic and readability constraints taking the cartographic symbols into account. When the goal is to derive a new database, it is called model generalisation and it is not constrained by cartographic symbols. Selection that can be considered as an operation concerning the abstraction of the database [Mackaness 2007] is a key step to model generalisation. It consists in choosing the relevant information in relation to the target map or database specifications. This paper presents a generic road network selection process based on data enrichment by spatial analysis.
The process has been tested on a large dataset extracted from BD Topo® data, a 1m resolution topographic database produced at IGN France. Several hypotheses are made on initial data : to be more generic, the attribute data used is poor (only road function and no traffic directions for example); other road attributes like road width or elevation can be useful but are not necessary; facilities point data has to be associated with the network to better understand it as roads are originally ways to communicate between facilities like housing areas, schools or industrial areas; finally, buildings or built-up areas are required to allow the discrimination of urban and rural areas in the network. The process has been tested to simulate selections with different target DLMs, thus using different set of parameters. One set of parameters corresponds to BD Carto®, a 10m resolution database produced at IGN. It allows a comparison between the selection and the actual database. The process presented in this paper has been implemented in Java on Clarity™ GIS from 1Spatial (former LaserScan).

Selection of the essential and relevant elements of a geo-database is more than choosing the important features and reducing the level of detail. As part of generalisation, it also aims at maintaining the main characteristics of geographic information. As roads are important features of maps and geo-databases, road network selection is a key topic of generalisation that has already been tackled in the past. Among others, graph theory techniques appeared to be very useful. They were introduced by [Mackaness & Beard 93] and were applied to road network selection in [Thomson & Richardson 95]. Moreover, some papers focused on street

networks, i.e. road networks in cities, where road density is very high and selection completely different [Ruas 1999], [Edwardes 2000], [Jiang&Claramunt 2004].
In order to maintain the characteristics of initial information, implicit information has first to be discovered using spatial analysis or pattern recognition [Mackaness & Edwards 2002]. Roads are complex man-made networks full of typical structures and patterns like roundabouts or grid streets networks as in Manhattan for example. [Marshall 2005] particularly focused on the patterns in street networks. Other papers deal with structure and pattern recognition in order to ease road selection [Thomson & Richardson 1999], [Heinzle et al 2005], [Heinzle et al 2006].

The next part of the paper describes why structure and patterns are essential in road networks and some new and existing methods for their recognition. Third part deals with the proposed selection process and particularly two main steps : rural and urban road selection. These two sections are illustrated with results obtained with a real database. Section four discusses the process and the results. Finally, last section draws some conclusions and discusses further research.

2. Data Enrichment by Pattern Recognition

Structure and pattern recognition is an essential initial step of generalisation [Mackaness & Edwards 2002]. Patterns can be defined as a property (shape, orientation, density...) within an object or between objects that is repeated regularly [Zhang 2004]. A structure can be defined as a particular distribution of objects representing an implicit geographic entity. Road networks perfectly illustrate this as roads form complex man-made networks full of particular structures and patterns. The detection of these structures is essential to maintain or transform them properly, as selection aims at reducing the level of detail while keeping the important and structuring information of the initial data.

Road crossroads can be considered as the atomic element of road network and it is very useful to better characterise them for most of next processes [Heinzle 2005]. The following paragraphs deal with the enrichment methods developed in order to allow the selection process presented in the paper. A simple taxonomy of crossroads has been designed based on [Grosso 2004]. The nodes of the network graph are classified as simple crossroads with T-nodes, fork nodes (Figure 1(b)) or star nodes for example. Such simple crossroads are detected, classified and then characterised by the geometrical and topological configuration of connected roads. For example, T-nodes (Figure 1(a)) are geometrically characterised by a nearly right angle α, a flat angle β and the absolute orientation of the crossroad. Topologically, it is characterised by a minor road leading to a major road.
Complex crossroads like roundabouts(Figure 1(e)) or highway interchange (Figure 1(d)) are particular aggregation of simple ones. For example, highway interchange may be characterised by clusters of fork and y-nodes in non built areas. Another example of useful complex crossroad is an offset crossroad, that is to say two opposite and very close T-nodes leading to the same major road. Offset crossroads may be key information in navigation purposes or in order to ease the caricature the offset.

Cul-de-sac (or dead ends) roads are another road structure necessary to detect as they are essential for the choice of important roads when linked to facilities access. Indeed, a dead end leading to an important facility play a key role in the road network. Obviously, facilities data has to be attached to the network to assess the accessibility role of cul-de-sac roads. Cul-de-sac roads and access are detected using road graph cycles.

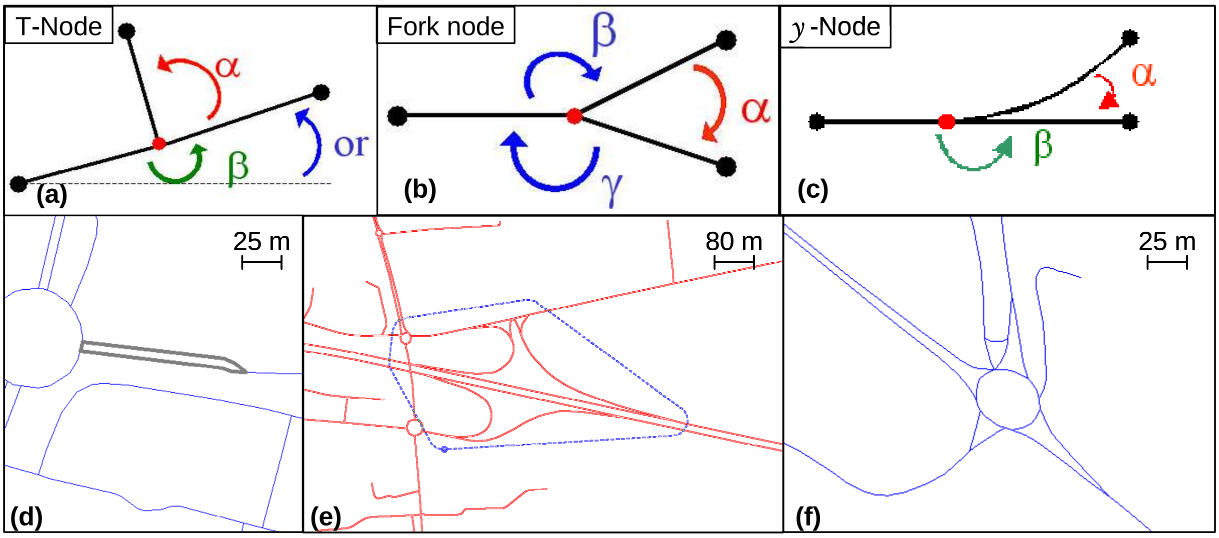

Figure 1 : (abc) : Three types of special simple crossroad. (d) : a branching crossroad. (e) : highway interchange detected using fork and y-nodes. (f) : a complex roundabout with branching crossroads. All the simple and complex crossroads are automatically detected.

Strokes, a structure introduced by [Thomson & Richardson 1999], are also created in our process. The notion of stroke is based on one of the perceptual grouping principles enunciated by the Gestalt Psychologists [Wertheimer 1938], namely the "good continuation" principle. Strokes are group of roads gathered by continuous curvature (Figure 2). Here, strokes are built taking into account road function (derived from attribute values) continuity added to curvature in order to be more relevant patterns. Besides, the stroke creation algorithm is modified to manage previously detected complex junctions such as roundabouts or branching crossroads : the strokes would be stopped at such junctions using only curvature.

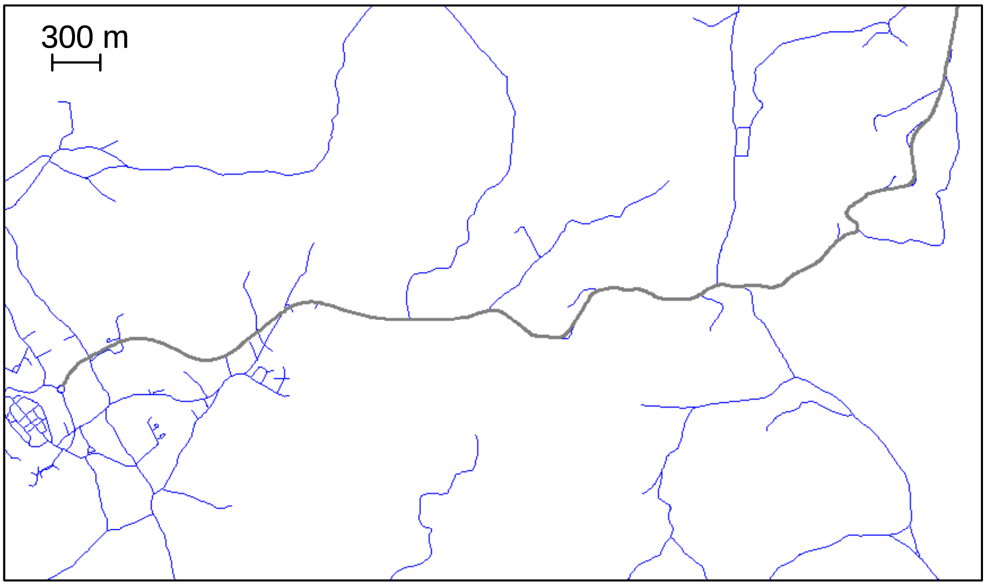

Figure 2 : a stroke in a road network : a group a roads following the "good continuation" principle.

Finally, more complex road patterns are recognised as their particularity requires adapted selection processes. First, dual carriageways (Figure 4(a)) detected using strokes and traffic direction in [Thom 2005] are in the process characterised only with geometry as traffic direction information is absent. Thus, road blocks, the faces of the network graph, are classified by three measures, convexity (Figure 3(a)), elongation (Figure 3(c)) and compactness (Figure 3(b)) that help to determine whether the face is thin (it is a dual carriageway) or not. Whether the polygon is convex or concave, elongation or compactness is preferred.

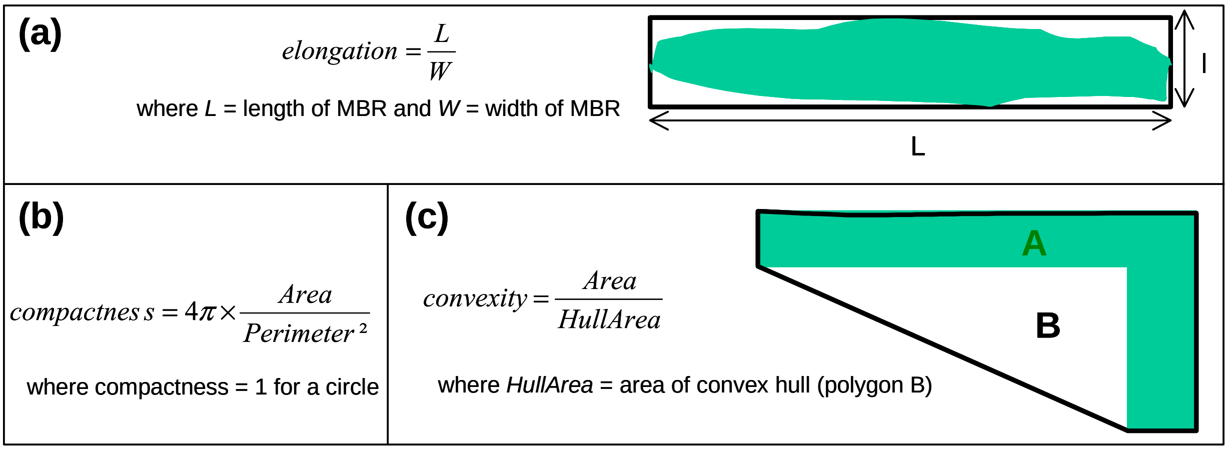

Figure 3 : the three measures used to evaluate the thinness of a polygon.

If the information is not in the initial database, roads belonging to highway services can be recognised using Y-nodes, topology (connexions with main roads) and the geometry of the road blocks (Figure 4(b)). Added to that, particular housing estate structures that would require rather typification than selection like comb or ladder structures [Heinzle et al 2005] can be detected using T-nodes (Figure 4(d)). Finally, grid structures that are important patterns of some particular networks (for example US cities) that should be maintained by selection can be detected using an algorithm presented in [Heinzle & Anders 2007]. The result of the automatic detection of such grids is shown in Figure 4(c).

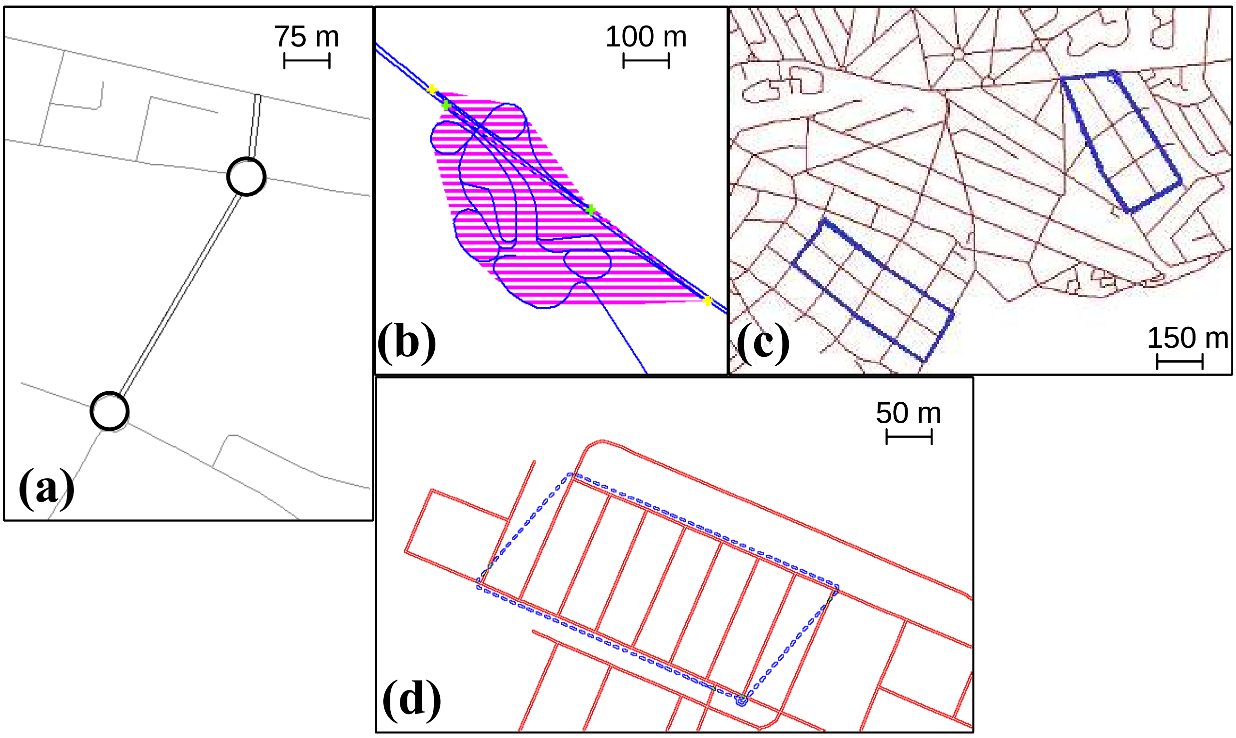

Figure 4 : (a) dual carriageways and roundabouts detected by automatic process. (b) : highway service or rest area structure automatically detected. (c) grid structures automatically detected. (d) a ladder housing estate road pattern automatically detected.

3. The selection process

The previous section presents a prerequisite step for the selection process described in this section as all the structures and patterns recognised in the road network enable a correct and relevant selection. But road networks are very complex and heterogeneous among countries and even areas of same countries. Rural and urban areas have completely different kind of network structures where road density, length or function may vary a lot. Hence, it does not seem appropriate to try to design a selection process that would deal with all kind of areas. Thus, two different selection processes were developed, one for rural areas, one for urban areas. Added to that, a continuity and consistency checking process at the interface of the areas and a structure typification step are included (Figure 5). Each step will be described in the following sections. The differentiation between rural and urban areas requires buildings or built-up areas in the initial database. Here, urban areas are built with buildings using the methods presented in [Boffet 2000].

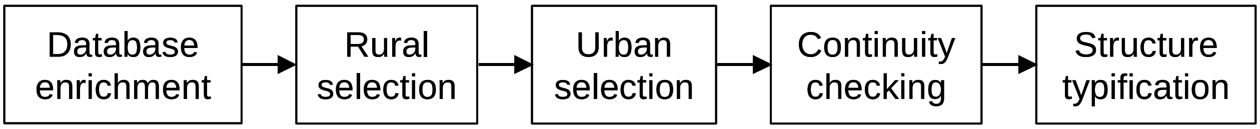

Figure 5 : Overview of the global road net work selection process

3.1 Road selection in rural areas

In rural areas, important roads are the ones that allow to travel from town to town. A common way to simulate this road attendance is the use of shortest paths in the network graph between attraction points of the graph [Richardson & Thomson 1996]. Rural selection is in this process also based on determining roads saliency by shortest paths. Globally, the important roads will be the ones often used by shortest paths. Shortest paths are computed in the whole road graph including urban areas to be really meaningful and to be used also later in urban areas.
Instead of using simply road length to compute shortest paths, a travel time estimation is used. The travel time is estimated with a cost function (Figure 6) taking into account attribute values or geometric characteristics : road length, road function (main roads are designed to go faster), road width (or number of ways), altitude difference, road sinuosity [Mustière 2001], town crossing.

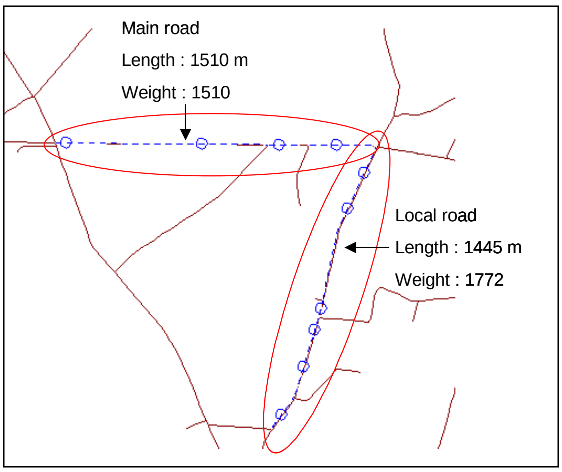

Figure 6 : example of cost function estimating road travel time

Shortest path is computed between attraction points of the network graph. Attraction points represent places where people go by car or that people leave by car. They can be housing estate area centres, schools, hospitals, commercial or industrial areas centres. In rural areas, attraction points can simply be villages centres to simulate moves from a village to another. Attraction points must be chosen in potential data of the initial database (i.e. schools can't be used when not present in the database) and they are weighted owing to attraction importance. Weights are given interactively, according to the type of attraction point, as parameters of the process. For example, airports generate much more traffic than a stadium and can be all weighted more. If the hypotheses on data concerning facilities cannot be fulfilled, other methods can be used : for example, non-weighted attraction points can be chosen randomly among the nodes of the road graph like in [Ruas & Morisset 97]. Then, shortest paths are computed using Dijkstra algorithm [Dijkstra 1959].
Like in [Richardson & Thomson 1996], roads used by shortest paths are given a traffic estimation value to assess their importance in relation to initial and final attraction point weights (see Figure 7). Figure 7(b) shows how roads are valued by the shortest paths from the circled attraction point to all others. For example, what is the contribution of the path from the circled initial attraction point to the two-weighted bottom right attraction point? The final attraction point has a weight of 2 so the value of all roads used by this shortest path are incremented by 2. Figure 7(c) shows the values of all roads after all shortest paths have been computed.

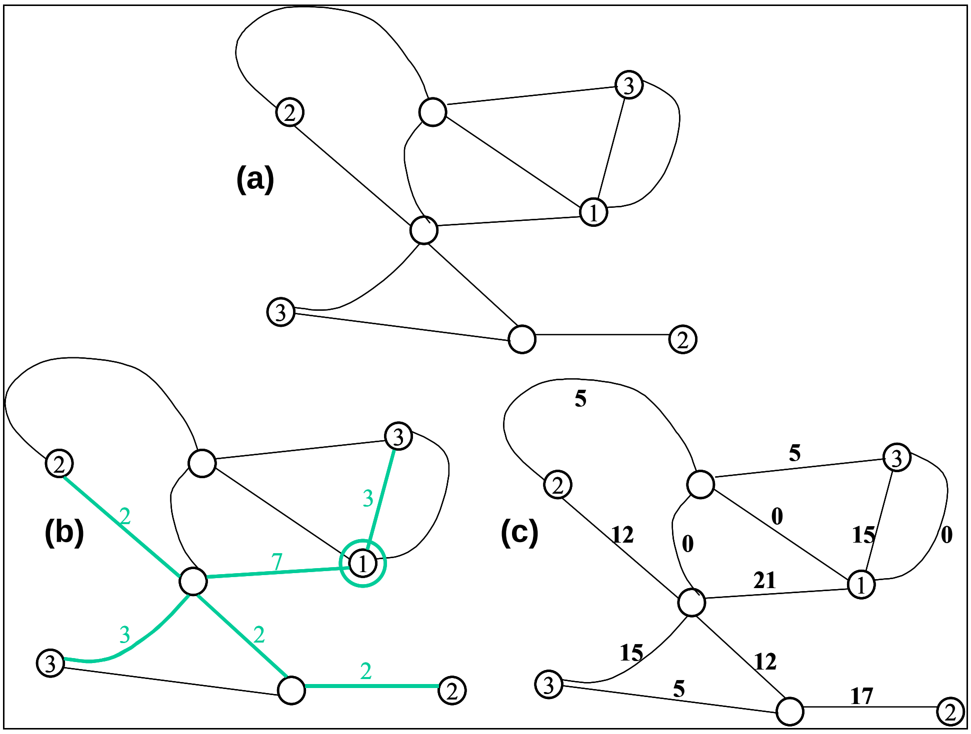

Figure 7 : valuing of roads in shortest path computation. In picture (a), there is a graph with 5 attraction points weighted from 1 to 3. In picture (b), the shortest paths from circled attraction point to the others are computed and road values are incremented as soon as they are used by a shortest path. In picture (c), once all shortest paths are computed, roads have their final values.

Finally, if the process is carried out on the entire dataset including urban areas, only the roads belonging to rural areas are selected according to the results obtained by this process. Selection is processed using different parameters that correspond to thresholds. A road is selected if :

- estimated road attendance is bigger than a given threshold.
- its corresponding stroke length is bigger than a given threshold.
- it belongs to a cul-de-sac longer than a given threshold.
- it belongs to a cul-de-sac with facility access longer than a given threshold.
- it is not part of a previously detected structure (structures are handled later in the process).

The process has been applied with parameters corresponding to BD CARTO® specifications (10m resolution database). **Figure 8** (a), (b) and (c) show some results of rural selection compared to the same extract in BD CARTO®. The results are quite similar and equivalent. An equivalent selection consists in selecting exactly the same important features and the same quantity of features but accepting small differences in secondary objects. As the target and initial databases may be totally independent in their management and their creation, up-to-dateness differences and biases in fuzzy survey specifications forbid to select exactly the same real world objects as in target database. **Figure 8**(d) and (e) shows rural selection with different parameters for a less detailed database than BD CARTO®.

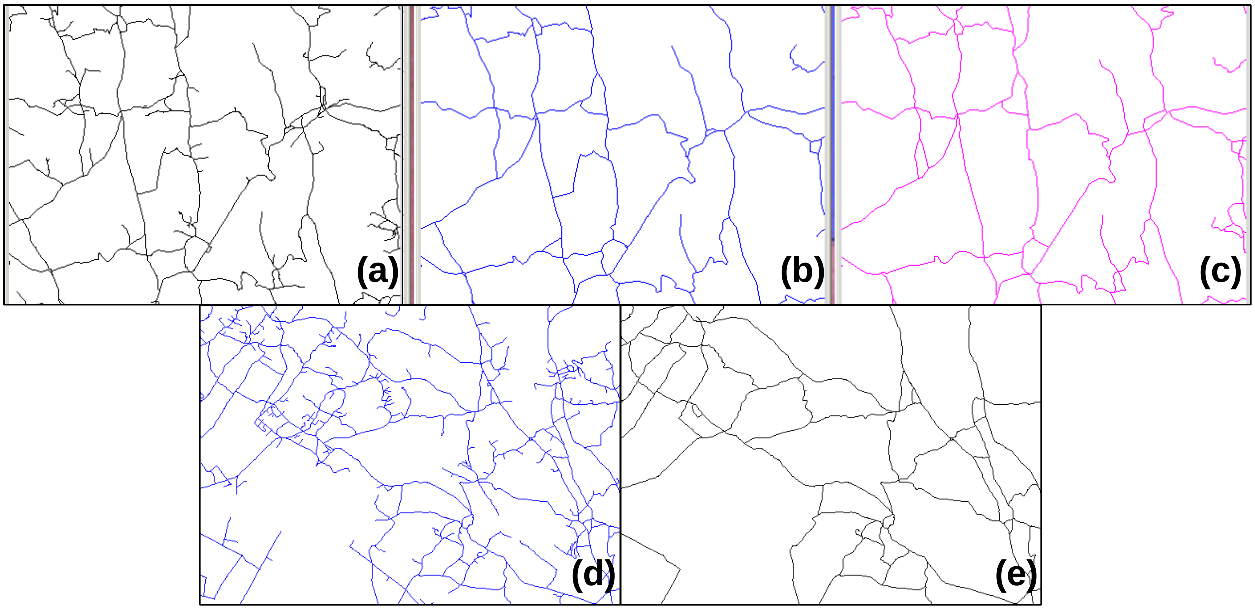

Figure 8 : the 3 first pictures (a, b and c) represent the same rural extract of a road network. (a) is the initial dataset (BD TOPO®). (b) is a selected dataset using the proposed process with the parameters of BD CARTO® specifications. (c) is actual BD CARTO®. (d) and (e) show a differently parameterised selection on another rural extract of the network.

3.2 Street selection in urban areas

3.2.1 Urban network enriched model

As street selection appears to be a more complex problem than rural road selection, different approaches have been tried [Ruas 1999], [Edwardes 2000] or [Jiang & Claramunt 2004]. The approach presented in this paper is a kind of synthesis of these three ones based on the road block aggregation principle of [Ruas 1999]. [Ruas 1999] aimed at cartographic generalisation at medium scale so buildings and buildings density were the keys for aggregation. Hence, the principle was kept but key factors for aggregation and dynamics were changed.

In our process, data has to be enriched in urban areas (here called "towns" for the sake of simplicity (Figure 9)). Towns are partitioned by road blocks (the faces of the road graph) called "urban blocks" that are to be aggregated in the selection process. Then main roads passing inside the town, determined with stroke length, road attribute function and traffic estimation (result of rural selection), are classified as urban axis and will not be eliminated in the process. Finally, towns are partitioned by urban axes to create "urban partitions" where block density is homogenous: road block density in the partition parameterises the level of block aggregation in the partition. It allows to maintain road density differences like in town centre (many small roads) compared to suburbs (few longer roads). Towns and partitions are characterised by measures like block density, stroke length mean or traffic estimation mean to allow consistent road block aggregation in different towns and partitions. All these new objects and measures that are presented in Figure 9 are computed automatically.

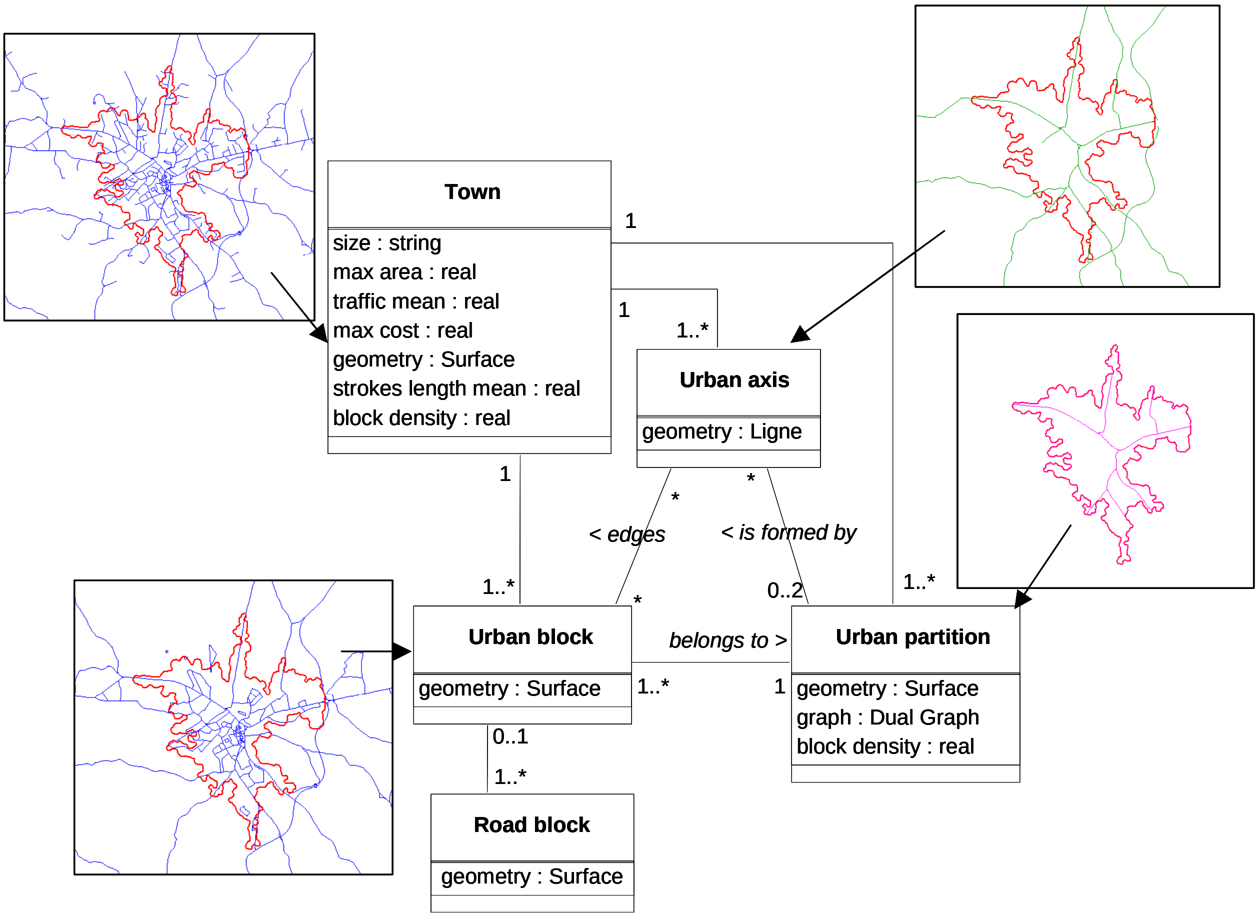

Figure 9 : enriched data model of an urban area for street selection, illustrated with data samples from the town of Salies-de-Béarn in the southwest of France.

3.2.2 Stroke centrality computation

Moreover, using the approach of [Jiang & Claramunt 2004], dual graphs of strokes are built for small towns and each partitions in large towns. In large towns, dual graphs of partitions are handled independently to save computation time as centrality computations are very long. In the dual graph, each stroke is a node and it is linked with the nodes representing all intersecting strokes. Centrality measures (degree, closeness and between-ness) are carried out on the dual graph nodes to assess the importance of these nodes and thus strokes importance in the street network. Degree, closeness and between-ness centrality are graph theory measures [Wasserman & Faust 1999] that evaluate the importance of the nodes in the graph (Figure 10). Degree centrality measures the number of connections of the nodes. Closeness centrality measures the length of the longest of the shortest paths between the node and all others : a node is central if it is close to all other nodes. The idea is that a node is central if it can quickly interact with all other nodes: in relation to road strokes, it is meaningful because a stroke is important if it can lead quickly to anywhere in the graph. Finally, between-ness centrality takes into account the fact that interactions between two nodes may depend on the centrality of the nodes located between them. In graphs, this centrality measures the use of a node in the shortest paths that connect all pairs of nodes. In this case, nodes that play a "bridge" role are very central as most shortest paths have to pass through the node (i.e. through the corresponding stroke in the road network). These three types of centrality express different kinds of importance for a node in a graph and so they are all used to assess strokes importance in a street network. Strokes role in the street graph is thus well characterised if the three centralities of the corresponding node is used. To avoid local biases, the centralities are divided by the mean of the corresponding centrality in the partition (or small town). Then, in order to avoid that one of the centralities influences more saliency and to simplify, a mean of the three centralities is computed.

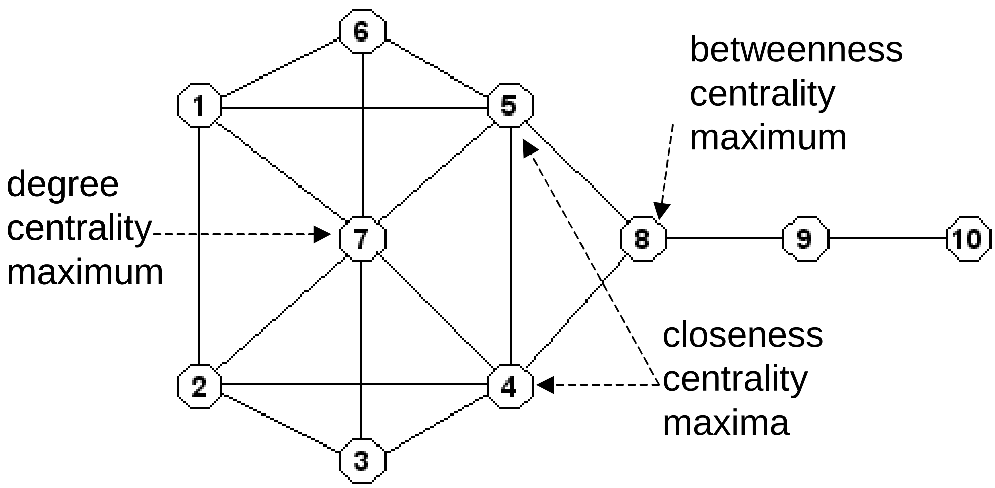

Figure 10 : In this graph, node 7 has the highest degree centrality, node 4 and 5 the highest closeness centrality and node 8 the highest between-ness centrality.

3.2.3 Dynamics of the aggregation

The aggregation algorithm principle is based on the computation of an aggregation cost. For each small block that needs to be aggregated, an aggregation cost is computed for each neighbour of the same partition (two blocks of different partitions can't be aggregated). The candidate neighbour with the lowest aggregation cost is chosen and aggregated if the cost is lower than a maximum cost. When there is no block small enough left or when all aggregation costs are above the maximum cost, the algorithm stops.
The cost function formula below takes several factors into account.

$$\cos t = (1 - compactness)^3 \times \sqrt{centralities\_mean} \times \frac{traffic\_estimation}{traffic\_estimation\_mean} \times \frac{stroke\_length}{stroke\_length\_mean} \times \sqrt{\frac{partition\_density}{town\_density}} \times size\_factor$$

The main factor is compactness computed with Miller's measure [Campbell 2000] (Figure 3(b)) as compact blocks are a good final result of the algorithm. The more appropriate the aggregation is, the lower the cost is. All factors used are around 1. If a factor is higher than one (for example the estimated traffic on the separating street is higher than the mean of the traffic estimations for the streets of the town), the cost increases to make the aggregation more difficult. The factors used are: the centralities, the traffic estimation of separating streets, the separating strokes length, the block density of the partition or the size of the aggregated block.

Once the algorithm has stopped, streets laying totally in an urban block are simply eliminated. Cul-de-sac selection in urban areas is managed the same way as in rural areas as it cannot be managed by a block aggregation process. The two parameters of the algorithm are the minimum size of a block and the maximum aggregation cost that is related to the cost function formula.
Figure 11 shows two street selections obtained using the aggregation process with two sets of parameters. The first case (b) consist in a light selection of the network while the second one is a stricter selection that approximately correspond to a 10m resolution database. In both, main streets still appear and the density difference between the city centre and the suburbs is visible. Such results are shown before the pattern typification step so roundabouts and dual carriageways (artefacts), visible in the pictures, will disappear according to the target specifications in the final results.

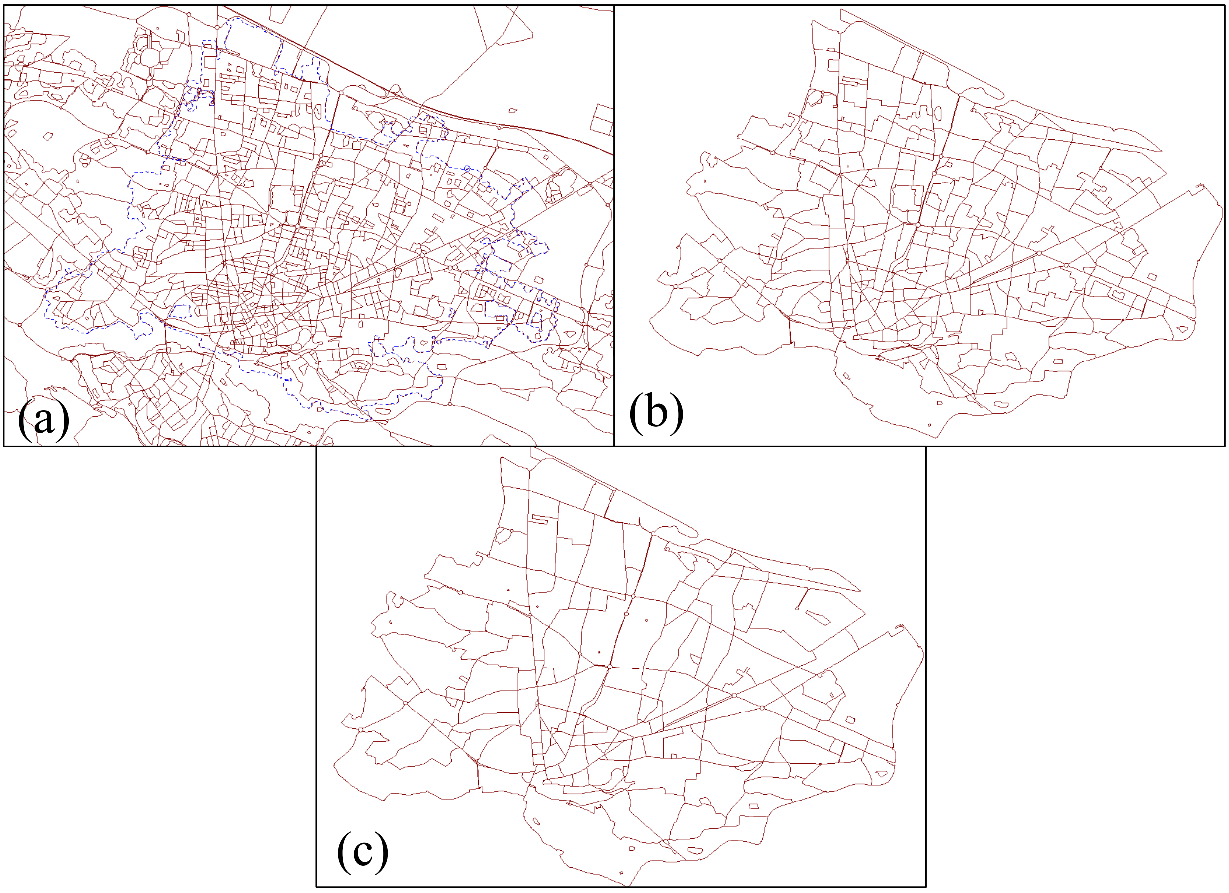

Figure 11 : (a) the street network of the city of Pau (southwest of France) in BD TOPO®. The extent of the urban area is in dashed blue. (b) a light selection of this street network. (c) a stronger selection corresponding to 10 m resolution database.

3.3 Continuity checking

The method for road network selection presented in this paper uses two different kind of processes for rural and urban areas. It would cause discontinuities and inconsistencies if they were used independently. In order to avoid such problems, a continuity checking step is necessary, particularly to control urban/rural interface zones.

The idea is that a road partly selected in one of the processes should have its other part selected. If it has not been selected by the other process, it is restored during the continuity checking step. Strokes seem to be the ideal way to allow this (Figure 12). In this example, the road is only selected in urban selection and continuity checking allows to select the part of the stroke that was not selected providing a more logical selected network.

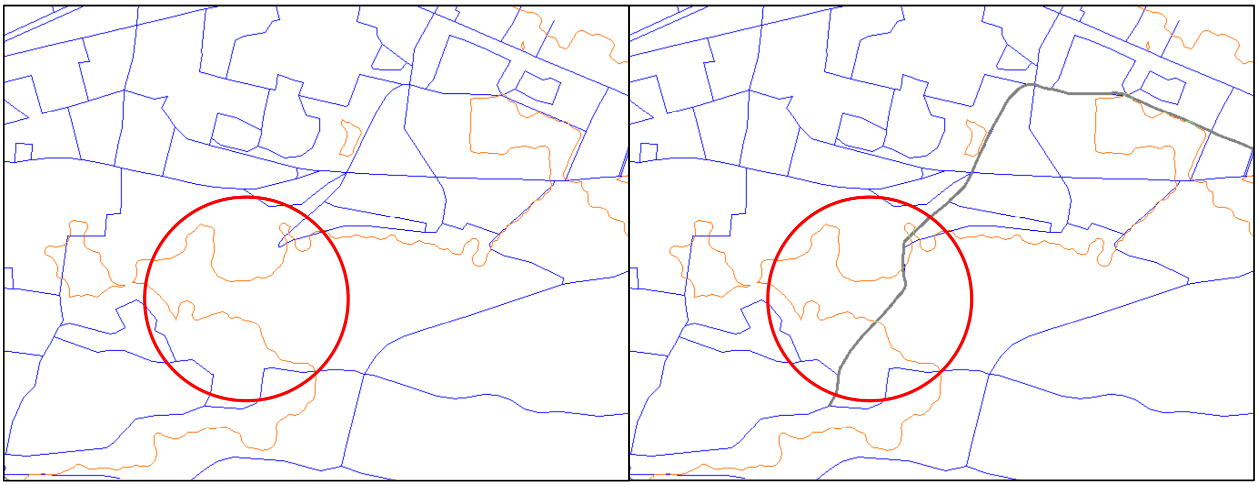

Figure 12 : On the left, the network is disconnected after the selection because of urban area limits. On the right, as one part of the stroke has been selected during urban selection step, the other part in the circle is also selected thanks to continuity checking.

Furthermore, to avoid discontinuities that are internal to one of the areas (it is most likely to appear in rural areas), an additional continuity checking post-process is carried out: roads that are not connected to the rest of the network are detected and the roads used by the shortest way to connect them are added to selection.

3.4 Patterns typification

Pattern typification or structuration can be considered as the final step of the selection process. Typification operator has been defined by [McMaster & Shea 1992] like this : "A selective number and pattern of the symbols are depicted. Generally, this is accomplished by leaving out the smallest features, or those which add little to the general impression of the distribution (...)or using a representative pattern of the symbols...". The process is not really selection as "while comparing the original with the derived map, it does not have to be obvious, which objects turned into the new ones." [AGENT 1999]. But it should be included in the method. This step concerns most road patterns that have been recognised in data enrichment step of the process (section 2). The road segments belonging to a pattern or structure have been left apart by the previous selection steps (rural, urban and continuity checking) in order to be finally typified. As a consequence, a new parameter of the overall process is the choice of the important road patterns that have to be typified rather than to be simply selected or eliminated by the appropriate process (rural or urban selection). For example, a specification of the target database BD CARTO® is : "road segments that form highway services or rest areas must be omitted and replace by a polygon object of class Highway service". In this example, highway services patterns have to be taken into account and to be transformed into a polygon, which is a kind of typification.

First, a simple structuration of roundabouts, branching crossroads or dual carriageways can be the collapse of the structure depending on thresholds. For example, specifications can consider the modelling of a roundabout as follows: it is a point if its diameter is under 100m and it is represented by ring roads else. The collapse of dual carriageways and branching crossroads can be done when needed using the straight skeleton algorithm [Haunert & Sester

2004] or Delaunay triangulation like in [Thom 2005]. The latter method was implemented (Figure 13).

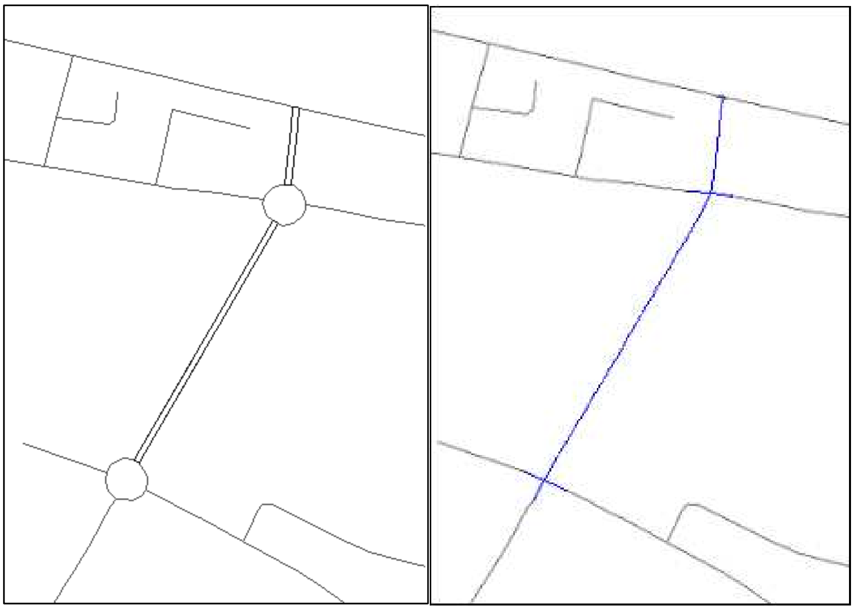

Figure 13 : exemple of structure typification, the collapse of dual carriageways or roundabouts (on the right, the collapsed geometry).

Regarding interchange structuration, it is a more complex problem that has been tackled partly by [Mackaness & Mackechnie 1999] that dealt more generally with road junctions. Adapted to interchange junctions and coupled with dual carriageways collapse, the algorithm provides interesting results (Figure 14). Order 1 is the most drastic degree of typification and shows interesting results provided dual carriageway is collapsed and geometries smoothed. Order 3, a slighter typification, is also an interesting simplification of the interchange.

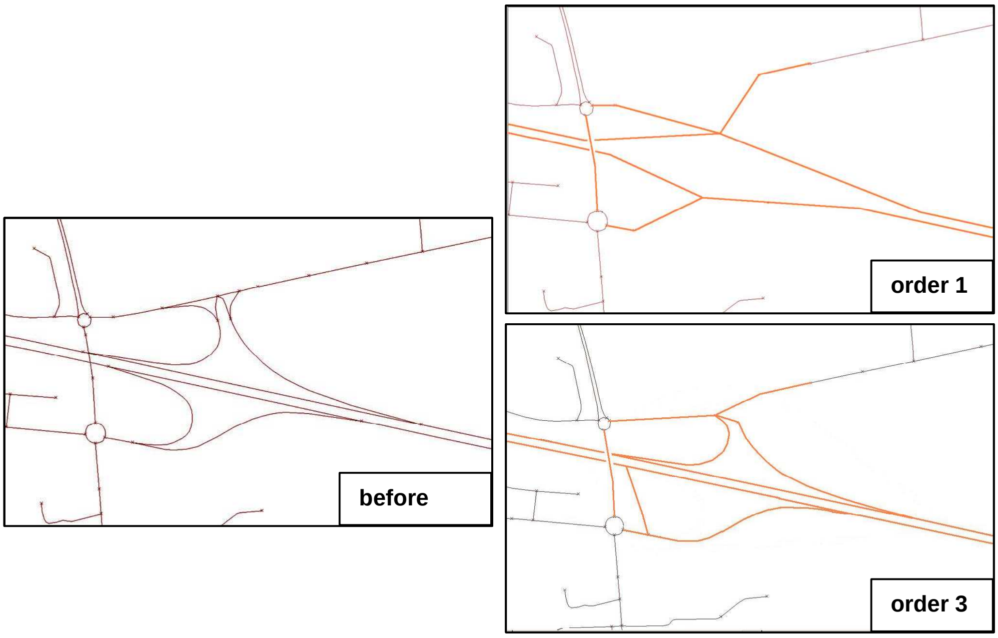

Figure 14 : interchange typification : on the left, before generalisation and on the right, 2 levels of typification with different parameters. In both cases, dual carriageways collapse is needed to finish the typification.

Regarding the other typifications needed in relation to the structures detected (see section 2), they have not been implemented yet but methods can be proposed. For example, highway services (Figure 15) can be typified by joining the enter and the exit road segments with a straight road segment and to omit all other roads.

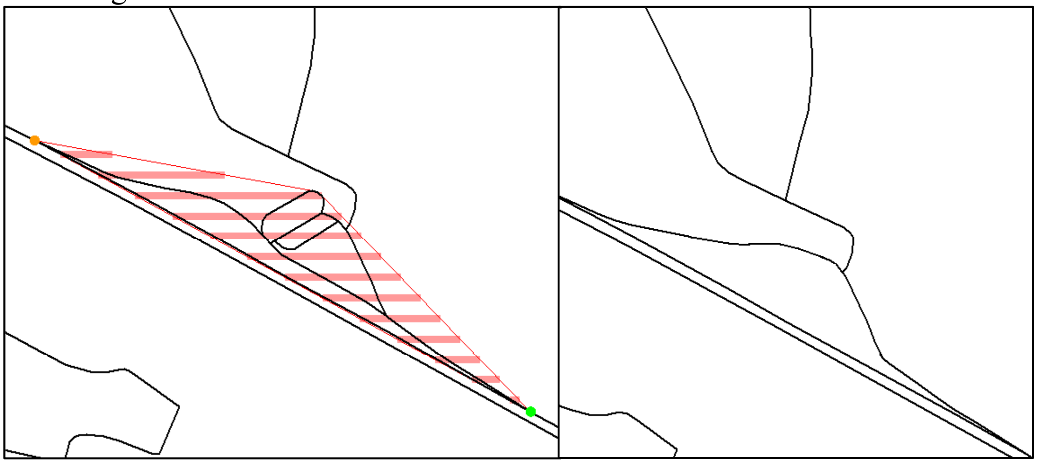

Figure 15 : On the left, an automatically detected rest-area and on the right, a typified version of the rest-area (obtained manually).

4. Discussion

The previous sections of this paper present the different steps of an overall process: database enrichment, rural selection, urban selection, continuity checking and structure typification. Except typification in some cases, each step is completely automated which allows it to be integrated to a more global automatic generalisation system. Moreover, the generic linking of these steps is also a contribution of this paper. Indeed, if each of the steps can be modified to be applied to particular road networks (more or less patterns, modification of block aggregation cost criteria in urban selection for example), the choice and the linking of these steps seem generic and useful in any case.

The results obtained with the implemented process on the Clarity platform can be discussed. The difficulty is that the effectiveness of a selection is hardly quantifiable because of the fuzziness of most selection specifications. Even when the result is compared with a reference dataset, it is hard to say if the differences are due to imperfect process, uncertainties in the translation of the specifications or errors in the reference target dataset (up-to-dateness for example). Nevertheless, the results were compared in our test case with BD CARTO®, a 10 m database managed and created independently from BD TOPO®, our initial database. Rural selection selects 95% of BD CARTO® roads (in terms of road length) and urban selection provides a road length up to 98% of BD CARTO® road length but not exactly the same road segments (75%). In this case, most differences are due to specification fuzziness and up-to-dateness differences.

In terms of efficiency, it is a quite heavy process with a lot of enrichment and process time consuming algorithms like shortest path and centrality computing, but it provides interesting results for database generalisation purposes.

The aim of this work was to design a generic process in relation to the target specifications but also to the initial data. Nevertheless, data specificities like attributes are sometimes used to make the process more effective. So, in order to be used with different datasets, the selection process would require little adjustments especially for shortest path computation

where additional data is used for node and link weighting (point facilities data, road attributes like width, function or elevation).
A difficult part of the process design is the choice of the multiple parameters and thresholds. Indeed, the translation of the target database specification may not always be obvious. When target specifications are too fuzzy to set the thresholds, they can be adjusted using Töpfer radical law [Töpfer & Pillewizer 1966] like in [Chaudhry & Mackaness 2005].

5. Conclusion and further work

To conclude, this paper presents a full and generic process to allow road network selection in model generalisations. It consists in (1) enriching the data by structures and patterns recognition, (2) rural selection based on assessing traffic by shortest path computing, (3) street selection algorithm based on road block aggregation and (4) structures typification. Tests carried out on large datasets show encouraging results for each step and for different kind of generalisation specifications.

To go further, pattern typification could be improved using for example [Anders 2006] for grid typification or developing and implementing better algorithms for interchanges, rest areas or housing estate structures. Interchange typification also has to be improved characterising more precisely highway ramps in order to typify them. Added to that, it would be interesting to integrate circular road detection to the process like in [Heinzle et al. 2006] to enhance the street selection algorithm. And the idea exposed in [Marshall 2005] that road patterns are like an unfolding fractal shows that many more structures can be detected in a road network to add knowledge and improve the selection.
As a further work, it would also be interesting to work on expressing the specifications of a road network selection in order to parameterise the process, like in [Hubert & Ruas 2003] that used generalised data samples to converge to the user needs. Finally, it could be interesting to insert this process into a global cartographic generalisation process that could thus deal with road selection before the cartographic operations.

References

AGENT project (1999), *Selection of Basic Algorithms***,** AGENT project technical report. Online documentation available at http://agent.ign.fr/deliverable/DD2.pdf
Anders, K.-H. (2006), *Grid Typification*. In proceedings of the 12th International Symposium of Spatial Data Handling'06 (SDH), University of Vienna, Austria, 2006.
Boffet A. (2000), *Creating urban information for cartographic generalisation*, In proceedings of the 10th International Symposium on Spatial Data Handling, (SDH'00), China, Beijing. 2000.
Campbell J. (2000), *Map Use and Analysis.* Dubuque : McGraw-Hill, 4th edition.
Chaudhry O., Mackaness W. (2005), *Rural and Urban Road Network Generalization Deriving 1:250,000 form OS MasterMap*, in proceedings of ICA2005, the International Cartography Association Conference, La Coruña (Spain), 2005.
Dijkstra, E. W. (1959), *A note on two problems in connexion with graphs*, Numerische Mathematik, 1, pp. 269-271.

Edwardes A.J., Mackaness W.A. (2000), *Intelligent Generalization of Urban Road Networks*. In proceedings of GIS Research UK Conference (GISRUK'2000), York, UK, 2000. pp. 81-85.
Grosso E. (2004). Etude des carrefours d'un réseau routier – Première approche : modélisation et implémentation d'outils de détection des carrefours (plate-forme Oxygène). Master thesis. COGIT Laboratory, Institut Géographique National. 2004.
Haunert J.-H., Sester M. (2004), *Using the Straight Skeleton for Generalisation in a Multiple Representation Environment.* In proceedings of the 6th CA Workshop on progress in automated map generalisation, Leicester, 2004.
Available from http://aci.ign.fr/Leicester/paper/Haunert-v2-ICAWorkshop.pdf.
Heinzle F., Anders K.-H., Sester M. (2005), *Graph Based Approaches for Recognition of Patterns and Implicit Information in Road Networks*, in proceedings of ICA2005, the International Cartography Association Conference, La Coruña (Spain), 2005.
Heinzle F., Anders K.-H., Sester M. (2006), *Pattern Recognition in Road Networks on the Example of Circular Road Detection*. in proceedings of the 4th International Conference GIScience 2006, Münster, Germany, September 2006. Full paper, pp 153-167.
Heinzle F., Anders K.-H. (2007), *Characterising Space via Pattern Recognition Techniques: Identifying Patterns in Road Networks.* Mackaness W., Ruas A., Sarjakoski T. (eds) : The Generalisation of Geographic Information : Models and Applications. Elsevier (2007).
Hubert F., Ruas A. (2003), *A method based on samples to capture user needs for generalisation*, 5th Workshop on Progress in Automated Map Generalization, Paris, France, 28-30 April 2003.
Jiang B., Claramunt C. (2004), *A Structural Approach to the Model Generalization of an Urban Street Network*, GeoInformatica, Vol. 8, Issue 2, June 2004.
Mackaness W. (2007), *Understanding Geographic Space.* Mackaness W., Ruas A., Sarjakoski T. (eds) : The Generalisation of Geographic Information : Models and Applications. Elsevier (2007).
Mackaness, W., Beard, K. (1993), *Use of Graph Theory to Support Map Generalisation.* Cartography and Geographic Information System (CaGIS), 20(4), 1993, pp. 210-221.
Mackaness, W., Edwards, G. (2002), *The Importance of Modelling Pattern and Structure in Automated Map Generalisation*. In Proceedings of Joint Workshop on Multi-scale Representations of Spatial Data, Ottawa, Canada (2002)
Mackaness W., Mackechnie G.A. (1999), *Automating the Detection and Simplification of Junctions in Road Networks.* GeoInformatica, Vol. 3, Issue 2, 1999. pp. 185-200.
MacMaster, R.B., Shea, K.S. (1992), *Generalization in Digital Cartography*. Association of American Cartographers, Washington D.C. (1992)
Marshall, S. (2005), *Streets & Patterns*. Spon Press, Taylor & Francis Group, New York 2005.
Mustière, S. (2001), Apprentissage supervisé pour la généralisation cartographique. PhD Thesis (2001).
Richardson D.E., Thomson R.C. (1996), Integrating Thematic, Geometric and Topologic Information in the Generalization of Road Networks. Cartographica Volume 33 (1) 1996, pp. 75-83.
Reynes, J.-L. (1997), *Sélection du réseau routier*. Master thesis, University Paris VI, COGIT Laboratory, 1997.
Ruas A., Morisset B., (1997), *Simulation and Agent Modelling for road selection in generalisation*, 18th International Cartographic Conference of ICA, (ICC'97, Sweden / Stockholm), vol3, p1376-1380.
Ruas, A. (1999), Modèle de généralisation de données géographiques à base de contraintes et d'autonomie. PhD Thesis (1999).

Thom S. (2005), A Strategy for Collapsing OS Integrated Transport Network Dual Carriageways, in proceedings of the 8th ICA Workshop on Generalization and Multiple Representation, La Coruña (Spain), 2005.
Thomson R., Richardson D. (1995), *A Graph Theory Approach to Road Network Generalisation*, in proceedings of the 17th ICC. ICA, Barcelona, Spain (1995). pp. 1871-1880.
Thomson R., Richardson D. (1999), The "Good Continuation" principle of Perceptual Organization applied to the Generalization of Road Networks, in proceedings of the 19th ICC. ICA, Ottawa, Canada (1999).
Töpfer, F., and Pillewizer, W. (1966), "*The principles of selection: a means of cartographic generalization*"; The Cartographic Journal, 3(1), pp. 10-16.
Wasserman S., Faust K. (1999), *Social Network Analysis : Methods and Applications.* Structural Analysis in the Social Sciences. Cambridge University Press, 1999. 825 p.
Wertheimer M. (1938). *Laws of organization in perceptual forms*. First published as Untersuchungen zur Lehre von der Gestalt II, in Psycologische Forschung, 4, 301-350. Translation published in Ellis, W. (1938). A source book of Gestalt psychology (pp. 71-88)., 1923.
Zhang Q. (2004), *Modelling Structure and Patterns in Road Network Generalization*, in proceedings of ICA Workshop on Generalisation and Multiple Representation - 20-21 August 2004 - Leicester