Correlation of Generalisation Effects with Thematic Attributes of Cartographic Objects

Yuen Hang CHOI and Zhilin LI

Department of Land Surveying and Geo-Informatics
The Hong Kong Polytechnic University
Hong Kong
Lszlli@polyu.edu.hk

ABSTRACT:

It has been widely recognised that, in order to automated map generalisation peocess, cartographic knowledge needs to be formalised. One way of knowledge acquisition is to make analysis of existing maps. Previous studies on this topic concentrated on the description of the phenomena on the surface. A typical example is the percentage change of the number of symbols or the percentage of open space on smaller scale maps. This study aims to go a step further, i.e to analyse the association of such changes with thematic attributes. In this study, road features on topographic maps of Hong Kong from 1:1,000 to 1:200,000 is studies; 6 types of thematic attributes, i.e. "type", "length", "width", "number of lanes", "number of traffic ways", and "connectivity", are considered; and, two statistic parameters, i.e. Lambda and Somers' Delta, are employed. The dependency of road elimination on these attributes is in the following order: type (0.73), length (0.53), number of lanes (0.48), number of traffic directions (0.41), width (0.36) and connectivity (0.19). If these numbers are normalised into percentage, then these values become: type (27%), length (20%), number of lanes (18%), number of traffic directions (15%), width (13%) and connectivity (7%). Such results could then be used to formulate an overall weight in determining determine whether a particular road should be deleted, merged or combined, in the generalisation process, in order to retain a certain percentage of road at a smaller scale.

1. Introduction

Maps are usually associated with scale. In each country or region, maps are normally produced in a series (with different scales). The maps at smaller scales are usually derived from maps at a larger scale to avoid the duplication of the expensive field (new) survey. The process to derive maps at a smaller scale from maps at a larger scale is referred to as map generalisation.

Maps need to be updated in a cycle of 2-3 years due to the changes in our environment. The traditional practice is to manually update maps at all scales and to maintain them. This is inefficient and expensive. A better solution, in this digital era, is to frequently update only the maps at the largest scale and then to derive maps at smaller scales by an automated generalisation process only when needed. Such an automated solution would be so beneficial economically and challenging scientifically that it has nowadays become a major international research theme in cartography and spatial information sciences. In this area, acquisition of cartographic knowledge is one of the major research topic (Weibel, 1995).

It has been recognised that knowledge of generalisation could be acquired from (a) cartographic experts through interview, (b) existing maps through analysis and (c) map specifications. A lot of work on his topic has been done (e.g. Buttenfield and McMaster, 1991). However, it has been generally agreed that the knowledge in cartographic experts' mind is not explicit enough to be extracted and the map specification normally tells "what not to do" instead of "what to do". Even there are specifications telling "what to do", they don't specify "how to do" in details. As a result, attention has been paid to study existing maps.

In other words, efforts have been made to find cartographic knowledge from existing maps. As topographic maps are the most standardised maps, they are normally selected for study. In this area, German cartographers have done some excellent work 30 years ago. Töpfer and Pillewizer (1966), through analysis of topographic maps, have formulated the so-called “Principle of Selection” or “Radical Laws”, expressing the relationship between map scale and number of features represented on maps. Their formulations are based on the implicit idea that generalisation is a process of “logical diminution”, which is widely accepted. Yu (1993) has extended this radical law by connecting it to the concept of "fractal dimension". A more comprehensive study of knowledge acquisition from existing maps has been carried out by Müller (1990). He has made a detailed analysis of the effect of generalisation on the number of features represented on German topographic maps at scales from 1:1,000 to 1:500,000. Not only the percentage number of features represented on these maps but also the ratio between the spaces occupied by different class of features were analysed. His results were represented in two different ways, i.e. logic representation scheme and charting the relationships between cartographic objects and generalisation operation into a relational table. Some rules have also been formulated. Leiter and Buttenfield (1995) have studied the topographic maps of Austria. They analysed the selection factors (percentage of features selected at a smaller scale) for carious features from 1:50,000, 1:200,000 to 1:500,000 scales. Some prototype rules have also been formulated.

However, in these studies, only the phenomena are described. The rules are still too general. A typical example is something like "If the scale changes from 1:50,000 to 1:200,000, then the number of individual settlement objects is reduced to 16% of the total at the larger scale". In such a rule, one still don't know which settlement to eliminate and which to combine. In order to answer these two questions, it seems that thematic (attribute) information could be a key. Indeed, the authors believe that rules based on thematic information should be formulated to control geometric operations and topological information should be used to validate the geometric operation and control the quality of generalisation process.

For reasons given above, the authors try to connect the changes of spatial representations at different scales to a set of thematic attributes.

2. Strategy of this study

In this study, the authors take the advice from Müller (1990), i.e. to "isolate the problems into modules, one for name generalisation, one for building generalisation, one for road network generalisation, independently from one another but in many possible different contexts. Only after, would their interaction be considered". As a result, only the transformation of the spatial representation of road network will be considered.

As topographic map has a long tradition and is regulated by more standards than any other map type (Müller, 1990), the series of topographic maps produced by local mapping agency, i.e. Survey and Mapping Office (SMO) of the Lands Department in the Hong Kong Government, is selected for study. In Hong Kong, topographic map series covers six scales, including 1:1,000, 1:5,000 and 1:20,000 1:50,000, 1:100,000 and 1:200,000. As it is not meaningful to study all maps covering the whole territory, a portion of Hong Kong Island is selected for this study because it is of best representatives of highly developed urban area as well as its hilly area with sparse houses. The map at 1:200,000 scale is shown in Figure 1.

Figure 1 Topographic map of the Hong Kong Island at 1:200,000 scale (from SMO)

The thematic attributes of road to be considered are 'type', 'length', 'width', 'number of lanes', 'number of traffic ways', and 'connectivity'.

The methodology of study consists of four stages:

(1) to observe and trace the changes in the representations of road features in different scale maps, i.e. to make an inventory;

(2) to quantify of the changes (transformations) into a value, so that it can be handled in the further statistical tests by use of "Feature Vanishing Level" (FVL);

(3) to carry out analysis of the dependency FVL to the these thematic attributes.

Statistical parameters will be used to represent the likelihood of feature with certain attributes to be eliminated (FVL) or preserved. Two statistical parameters, i.e Lambda (λ) and Somers' Delta (Δ) are used. More detailed discussion of these parameters will be discussed later in analysis section.

3. Observation of the transformations of spatial representation

For typical topographic maps at 1:1000 and 1:5000 scales, almost all map features are shown in their true dimensions. However, when the scale is further reduced, the resulting map features would be too densely packed and too small to be deciphered. One of the solutions to clearly present the map features is to symbolise the map features. At 1:20000, roads are classified in three types, i.e., expressway, main road and secondary road. Secondary road is further divided into five sub-classes, i.e., double width road (A), single width road (B), unsurfaced road (C), non-motorable road (D) and lane (E). At further smaller scales, a more general classification is formed. For example, expressway and main road are grouped together and represented by same symbol; double width secondary road and single width secondary road are reclassified as secondary road; while the unsurfaced secondary road is separated to form a new sub-class. However, this sub-class is eliminated when at 1:100,000 and 1:200,000. The "lane" and "non-motorable" roads are eliminated at 1:50,000. The transformation process could be illustrated in Figure 2. The change number of road segments during the scale reduction as shown in Table 1.

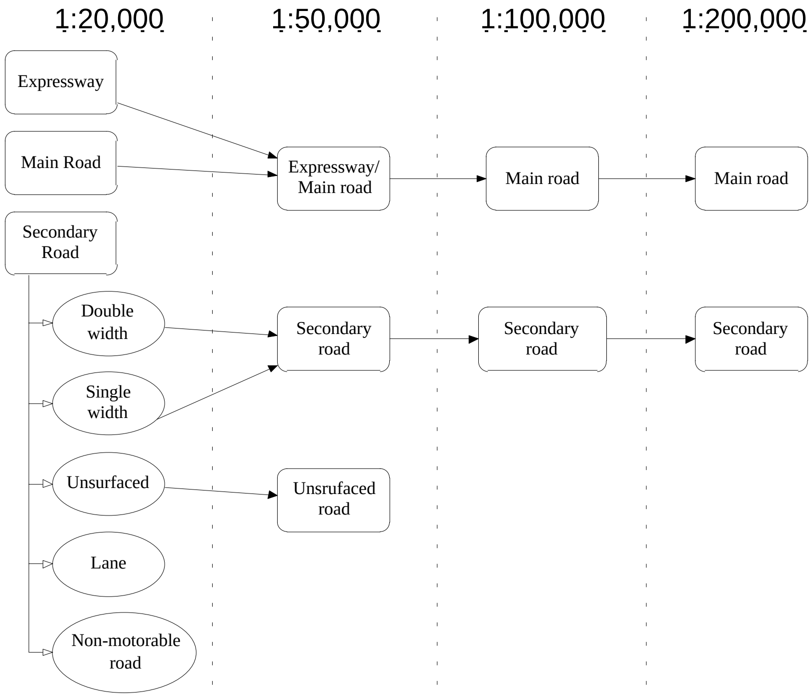

Figure 2 *Transformations of road features from 1:20,000 to 1:200,000*

However, things are not as simple as the observations described above. For example, only 25.4% of "secondary roads" are preserved from at 1:200,000. The question arising is "what kind of secondary roads is should be preserved and what kind eliminated and/or combined". There might be no answer to this question. Therefore, one would try to find a less definite answer, e.g. "Those secondary road with has connections in 4 or more directions are more likely to be preserved than those with two connections”. Then the question becomes something like "how likely would the secondary roads with 4 or more connecting directions will be preserved"? There are more similar questions. The next two sections will try to answer such questions.

Before going to discuss the quantification of changes of raod representation at different scales. A discussion of the classification systems needs to be made. There are two classification systems used by two government departments, i.e SMO and Transport Department. The difference somehow causes confusion on the study of transformations of road features with respect to its types. The classification of roads in the Transport Department is based on the function of roads and roads are classified into two networks, i.e. Major Road Network and Minor Road Network. The Major Road Network includes Expressway (EX), Urban Trunk Road (UT), Rural Trunk Road (RT), Primary Distributor (PD), District Distributor (DD), Rural Road A (RD), and some Local Distributor (LD). The remaining Local Distributors form the Minor Road Network. The classification is mainly based on their designed functions which are declared by their names and supplemented by the road construction standard and traffic management. The differences in classification between SMO and Transport Department can be found out by simple matching between these two classification schemes. However, there is no information about unsurfaced and non-motorable roads in the Transport Department. So these two classes cannot be compared, but can only be clarified by their definitions from the Lands Department. A comparison of the classification of roads between Lands Department and Transport Department reveals:

- All the Expressways are classified as ‘Expressways’.
- All the Urban Trunk Roads are classified as ‘Main Roads’.
- Nearly all Primary Distributor roads (96.6%) are classified as ‘Main Roads’. Only 1 Primary Distributor road is classified as ‘Secondary (type) A’ roads.
- About three-fourth of District Distributor roads (73.9%) are classified as ‘Main Roads’ and the remaining quarter of District Distributor roads (26.1%) are classified as ‘Secondary (type) A’ roads.
- A few Local Distributor roads (5.1%) are also classified as ‘Main Roads’. Most of Local Distributor roads are classified as ‘Secondary (type) A’ roads (84.8%). Some of them are classified as ‘Secondary (type) B’ roads (7.3%) and the remaining are classified as ‘Lanes’ (2.8%).
- Nearly all the Lanes (84.6%) are classified as ‘Lanes’, only a few (15.4%) are classified as ‘Secondary (type) A’ roads.

These two systems are quite consistent. In this study, the system by the transportation department is used as it system seems more meaningful thematically.

4. Quantification of transformations of road representations

Totally, 280 (at 1:20,000 or larger scale) road segments were observed from topographic maps of Hong Kong Island. In this section, changes in the representation of road feature at different scales will be quantified and statistical analysis will be carried out in the next section.

Road type serves as a major index of importance of different roads. During transformations of road representation from 1:1,000 to 1:200,000, not all less important roads are erased. At each scale, numbers of road segments grouped by road types were counted. The results are summarised in Table 2a and number in percentage is also showed in Table 2b. A graphic representation is the results in Table 2b is shown in Figure 3. It is clear that all UT and EX types of roads are retained. While the number of LD and DD types decrease more or less linearly as a decrease in scale from 1:20,000. All roads with LE type disappear from 1:50,000.

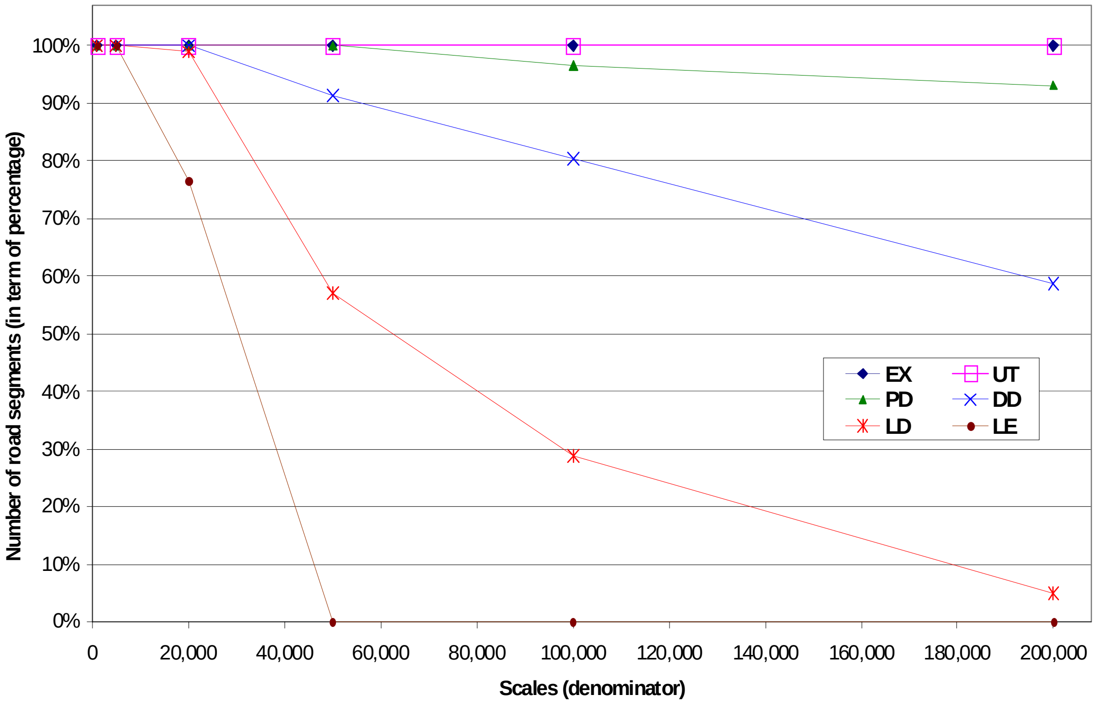

Figure 3 Retention of road segments (in percentage) for each road type at different scales.

By observation, it could be found that roads with different lengths may be erased or retained during scale reduction. There is no threshold to determine the elimination of road segments at each scale. To obtain quantitative observations, the total number of road segments on the maps at each scale is counted. For each scale, the road segments are divided into ten groups so that the first group occupies the first ten percentages, second group occupies the first twenty percentages and so on. Then the mean values of road length for each group are computed (Table 3). The results of a such categorisation is shown in Figure 4. It is clear that the reduction of short road segments is dramatic with a decrease in scale.

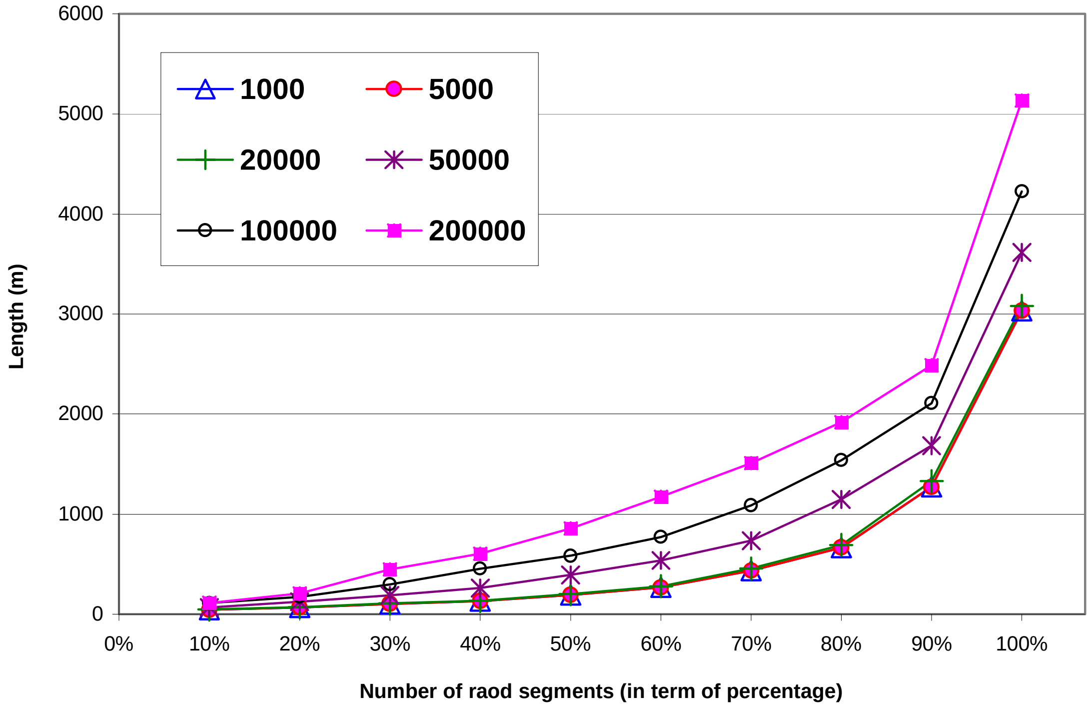

Figure 4 Mean road lengths at different portions of total number of roads for different scales.

In the case of road width, the situation is similar. Therefore, methodology used for the study of road width is the same as the study of road length. The means of road width with different percentage is shown in Table 4 and Figure 5. It appears that the mean widths of the first 30% road segments represented on maps at the scale 1:20,000, 1:50,000 and 1:100,000 are nearly the same. The occupying rate of wider road increase dramatically when the scale is 1:50,000 or smaller.

In the study of the association of road representation with number of lanes, those non-motorable road segment are excluded. There are 20 such road segments in total. The number of lanes for the remaining 260 roads ranges from 1 to 11. Table 5a recorded the total number of road segments which are shown at each scale for different number of lanes It can be found that, with fewer lanes, a road more likely they are eliminated. Elimination takes place mostly when the number of lanes is one or two. No roads with 6 or more lanes are eliminated. Table 5b shows the number of road segments (in percentage) at different scales for lanes from 1 to 5. Then the data are showed in Figure 6.

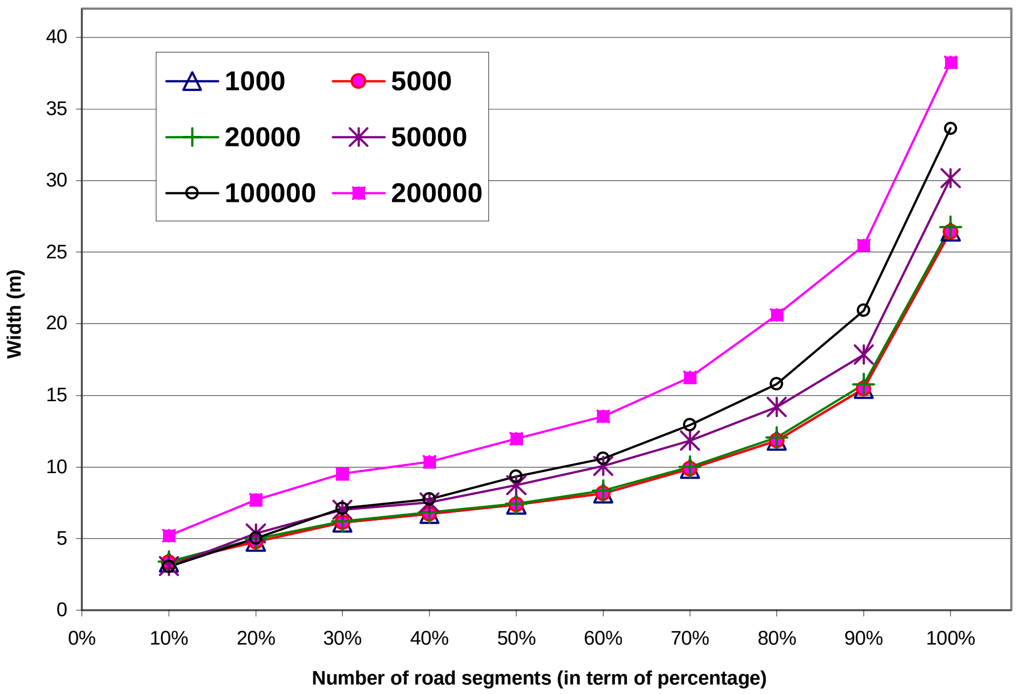

Figure 5 The relation between mean road widths and total number of roads for different scales.

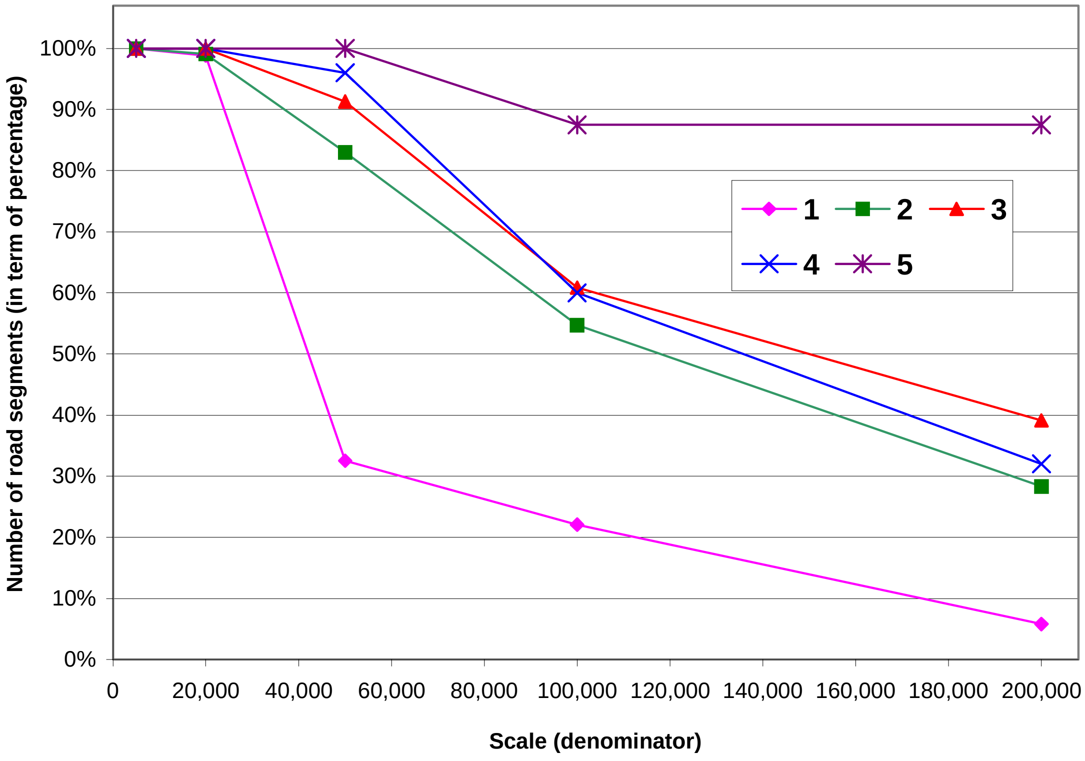

Figure 6 Retention of road segments (in terms of percentage) at different scales for lanes from 1 to 5.

A road is usually not designed for only one traffic direction. Most of roads are designed for bi-directional (i.e. transporting in reverse directions). However, some roads are non-motorable and hence the number of traffic ways is declared as ‘0’. Sometimes, a road is wide enough to serve for three traffic ways, so the number of traffic ways is declared as ‘3’. The categorisation of road segments by number of traffic ways is presented in Tables 6a and 6b. A graphic view is given in Figure 7. It is clear that almost all non-motorable roads disappear from 1:50,000. Road segments containing three traffic ways are kept at all scales. Road segments with more traffic ways are more likely to be shown on the maps

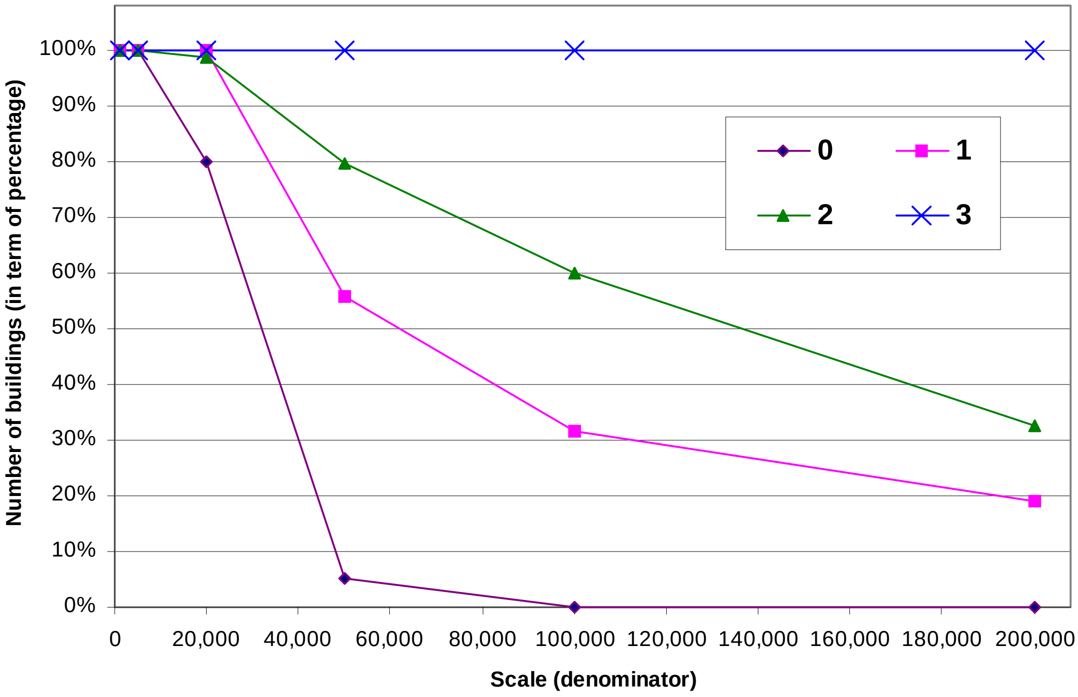

Figure 7 Retention of road segments (in terms of percentage) at different scales for traffic ways 0 to 3.

Connectivity is an important concept in network analysis. If a node is connected by two arcs, it is considered as having an order of 2. Similarly, if a node is connected by three arcs, it has an order of 3, and so forth. As a road segment consists of a node at each end, there is a need of a pair of two numbers to identify its order, e.g. ‘0,0’. The values of these pairs range from "0,0" to "5,5" in this study. The result of road segments categorised by this system is listed in Table 7a.

The data in Table 7a can be reclassified into five categories, ‘0,0’, ‘X,0’, '1,1', ‘X,1’ and ‘X,X’. ‘0,0’ means that the road segment is either a floating segment (without connection to other segments) or a self-contained loop. ‘X,0’ signifies a dangling segment with one end being unconnected and the other end being connected to any number of road segments, except number ‘0’. ‘1,1’ means that it is a bridge between two segments; ‘X,1’ also means a bridge but it is used to connect a segment and several segments at the two ends;

the ‘X’ in this case can be any number except 0 and 1. Finally, the ‘X,X’ means that the two ends can be connected to any number of road segments, but the number should be equal to or greater than 2. After the data are reclassified, new tables (Table 7b and 7c) containing the above five categories is formed and the relationship between connectivity and number of roads retained at different scales is illustrated in Figure 8. It is clear that those road with "0,0" connectivity are all eliminated from 1:50,000.

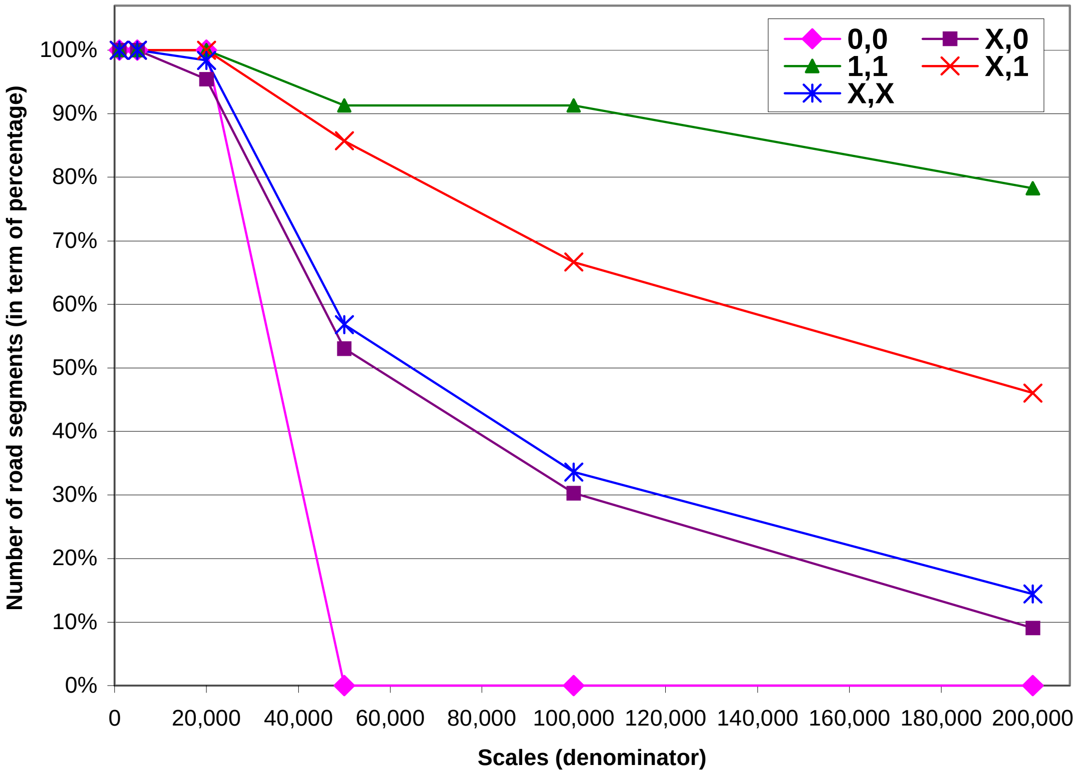

Figure 8 Elimination of road segments at different scales for each connectivity type.

5. Dependency of "feature vanishing level" to thematic attributes

The study quantified the strength and nature of the relationships between REL and six road attributes: type, length, width, number of lanes, number of traffic way, and connectivity.

Due to different nature of the attributes, data observed belong to two different scale of measurement, i.e. nominal and ratio. Two types of association (dependence) measures were applied, respectively. For those nominal data, the method used was Lambda (λ). For ratio data, Somers' Delta (Δ) is used instead.

Goodman-Kruskal’s Lambda statistic (see Norusis, 1988) is a proportional reduction in error (PRE) measure which looks at how much better to predict the values of a dependent variable when the values of an independent variable is known. It is expressed as follows:

$$\lambda = \frac{\text{Misclassified in situation 1 - Misclassified in situation 2}}{\text{Misclassified in situation 1}} \qquad (1)$$

A value of 0 for Lambda means that the independent variable is of no help in predicting the dependent variable. When two variables are statistically independent, value of Lambda is 0. However, a Lambda of 0 does not necessarily imply statistical independence. In this formula, errors for the situation 1 and situation 2 are explained by the examples shown by an example shown in Table 8. This example is used to predict whether life is seen as dull, routine, or exciting by using the information about the highest degree earned.

In this example, it is known through sample, that 5.1% people found life 'dull', 44.2% 'routine' and 50.8% 'exciting'. If one predicts 'Exciting' (which is of highest frequency) for everyone, the error for this prediction is the total number of misclassified, that is, 368 (38 + 330). However, if one's highest qualification is known, the error in prediction could be reduced. If one predict 'routine' for high school graduate and 'exciting' for both bachelor and higher degree holders, then the total error would be 266 (35 + 231) + 60 (2 + 58) and 22 (1 + 21). Comparing these two errors, the improvement is

$$\lambda = \frac{368 - 348}{368} = \frac{20}{368} = 0.0543 = 5.43\%$$

By knowing the highest degree a person has earned, the error is reduced by 5.4%.

Somers' Delta (Δ) is a kind of ordinal measure of association that is based on the difference between the number of concordant pairs and the number of discordant pairs. It is defined as follows:

$$\Delta = \frac{\text{Number of concordant pairs - Number of discordant pairs}}{\text{Sum of all pairs of cases that are not tied on the independent variable}} \qquad (5.2)$$

where, concordant pairs, discordant pairs and all other pairs are explained by an example .given in Table 9, which contains a cross tabulation of the values of ‘life’ (the view of life variable) and 'degree' (the education variable) for three cases.

In the situation of Case1 and Case 2, both values (Life and Degree) at Case 2 are larger than those at Case 1. Such a pair of cases is referred to as being 'Concordant'. In contrast, in the situation of Case 2 and Case 3, the value of 'Life' in case 3 is larger than that in Case 2 but the values for "Degree" is reversed. Such pair is called 'Discordant'. If the value for one variable is the same, the pair is referred to as being 'tied'. In summary, there are five possible combinations, i.e. concordant, discordant, tied on the first variable, tied on the second variable and tied on both variables. If most of the pairs are concordant, the association between the two variables is said to be positive. If concordant and discordant pairs are equally likely, it can be concluded that there is no association between these two variables. So '+1' indicates a perfect positive relationship; '-1' indicates a perfect negative relationship; and '0' indicates no relationship.

To carry out statistical analysis, the observed data are categories into six different ranked levels, i.e. 6 feature vanishing levels (FVL). As the features to be studied in this project is the road feature, a more specific term, i.e. road elimination levels (REL). Such a categorisation is as follows:

REL 1: Features eliminated at scale 1:5,000
REL 2: Features eliminated at scale 1:20,000
REL 3: Features eliminated at scale 1:50,000
REL 4: Features eliminated at scale 1:100,000
REL 5: Features eliminated at scale 1:200,000
REL 6: Features retained at all scale

Two statistical parameters are worthy particular attention. The first one is the value of ‘FVL Dependency’, indicating the strength of the association measurement. The larger the value, the stronger the association. The second one is the ‘Approximation Significance’ value, showing the necessity to reject the null hypothesis of the statistical tests, i.e. "FVL being independent of the tested variable. The smaller the value, the more strongly one should reject the null hypothesis.

Lets now turn to the association of FVL with road attributes. The relevant data are recorded in Tables 10 to 15. The Somers’ Δ and Lamda tests were carried out and results are showed in Table 16. It can be seen that the value of the approximate significance is less than 0.0005, so the null hypothesis, i.e. ‘FVL being independent of road attributes’, is strongly rejected. The strength of the association is indicated by either Dalta or Lamda values. In this test, the results show that all the associations are positive. But the value ranges from 0.733 to 0.193, from strong to weak. More precisely, the results are: The dependency of road elimination on these attributes is in the following order: type (0.73), length (0.53), number of lanes (0.48), number of traffic directions (0.41), width (0.36) and connectivity (0.19).

The result in graphic form is as shown in Figure 9a. If these numbers are normalised into percentage, then the result is something like Figure 9b. However, it might be too appropriate to do so as Δ and λ have different meanings.

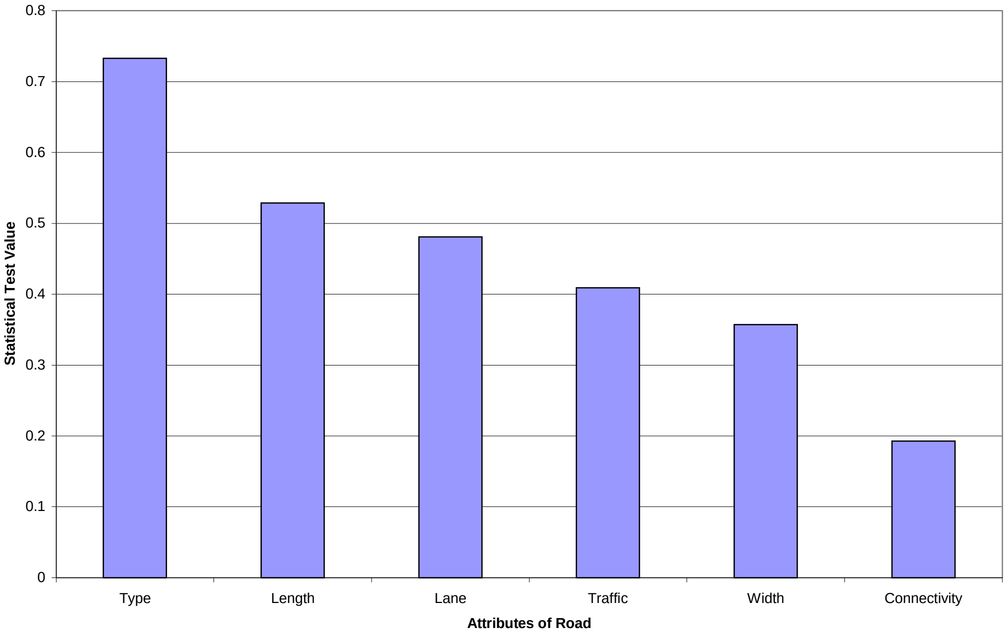

Figure 9a Dependency of the change in road representation on road attributes: Statistical results

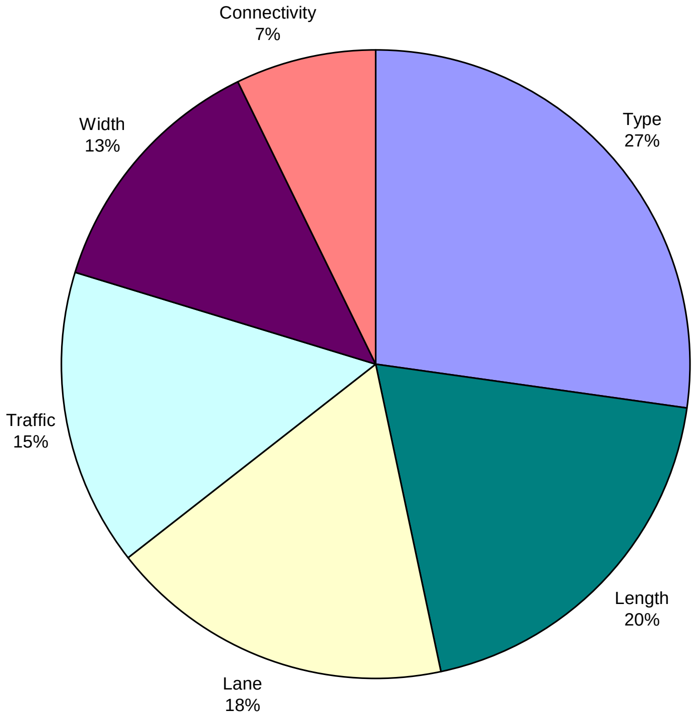

Figure 9b Dependency of the change in road representation on road attributes in terms of percentage

Discussion and conclusions

In the study, topographic maps six different scales, ranging from 1:1000 to 1:200,000, were selected for the study. Through observations and analyses, some empirical facts were found out about how roads were represented on maps at different scales. The change in the representation of road feature at different scale can be described as follows:

- In the 1:1000 and 1:5000 maps, roads are represented in their true dimensions by curb lines. The process of selective omission of road features is started at scale 1:50,000. The number of road segments remains unchanged until at scale 1:50,000, it has been reduced by 35% at that scale. Since then, the number of road features is decreasing, about 46% of road features are retained at scale 1:100,000 and only about 25% at scale 1:200,000.
- At 1:20,000, roads are classified and symbolised into five classes: Expressway, Main Road, Secondary Road (which includes five subclasses), Elevated Road and Road Under Construction.
- At 1:50,000, the Expressways and main roads are grouped together; the double width and single width secondary roads are classified as secondary road; the unsurfaced secondary roads are independently to form a new class; and the lanes as well as the non-motorable roads are erased.
- At 1:100,000 and 1:200,000, the unsurfaced secondary roads are erased and only the main roads and secondary roads are remained.

These are general statements. Their usefulness for rule formalisation is limited. To obtained for useful information, quantitative analysis is employed. In this study, the findings were quantified into several levels termed ‘Feature Vanishing Levels’ or FVL in short. Statistical analysis on the association of the change of the representation of road features with road attributes was conducted. 6 attributes were used, i.e. type, length, number of lanes, number of traffic directions, width and connectivity. Two statistical tests were employed, i.e. Somers’ Delta and Lambda.

The attribute ‘Road Type’ is the most important factor for the selection of road features to be retained or erased during scale reduction; whereas ‘Road Length’, ‘Number of Lanes’, ‘Number of Traffic Ways’ and ‘Road Width’ are in the similar significance; and, surprisingly, ‘Connectivity’ is the least important factor. As shown on Table 16, the value of Somers’ Delta of ‘Road Type’ is 0.733 which is outstanding (the greatest) from all the others, while the value of Goodman-Kruskal’s Lambda for ‘Connectivity’ is 0.193, and the values for others are from 0.3 to 0.5.

If these numbers are normalised into percentage, then these values become: type (27%), length (20%), number of lanes (18%), number of traffic directions (15%), width (13%) and connectivity (7%). Such results could then be used to formulate an overall weight in determining determine whether a particular road should be deleted, merged or combined, in the generalisation process, in order to retain a certain percentage of road at a smaller scale.

It must be noted that this is a first attempt to measure the dependency of changes in the representation of road features on road attributes. This work certainly has limitation. It is recommend that some other attributes may be considered and other statistical measures explored. The study should then be extended to other map features such as settlement.

Acknowledgment

Mr. Choi would like to thank the Hong Kong Polytechnic University for its studentship. The authors would like to thank Lands Department of the Hong Kong SAR government for providing the digital map data for this study.

References

Buttenfield B.P. and R.B. McMaster (Eds.) (1991). *Map Generalization: Making Rules for Knowledge Representation*, Longman Scientific & Technical, Essex, England.

Leitner, M. and B.P. Buttenfield (1995). “Acquisition of Procedural Cartographic Knowledge by Reverse Engineering.” *Cartography and Geographic Information Systems*, Vol. 22, No. 3, pp. 232-241.

Müller, J.C., 1990, Rule based generalisation: potentials and impediments, Proceedings of the 4th International Symposium on Spatial Data Handling, Zurich. 1:317-334

Norusis, M.J., 1998, SPSS 8.0 Guide to data analysis, Prentice-Hall, Inc.

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

Yu, Z., 1993, The effects of scale change on map structure, Doctor's Thesis, Department of Geography, Clark University

Weibel, R., 1995. Three essential building blocks for automated generalization. In: Müller, J.-C., Lagrange, J-P. and Weibel, R. (eds.), 1995. *GIS and Generalization*. Yaylor & Francis. 56-69.

Table 1 Change in number of road segments of different classes from 1:1,000 to 1:200,000

	1 : 1,000	1 : 5,000	1 : 20,000	1 : 50,000	1 : 100,000	1 : 200,000
Main Roads	--	--	79 (100%)	74 (93.7%)	74 (93.7%)	64 (81.0%)
Secondary Roads	--	--	195 (100%)	108 (55.4%)	56 (28.7%)	7 (3.6%)
Total	280	280	274 (100%)	182 (65.0%)	130 (46.4%)	71 (25.4%)

Table 2a Number of road segments at different scales in different types.

	1:1,000	1:5,000	1:20,000	1:50,000	1:100,000	1:200,000
EX	1	1	1	1	1	1
UT	7	7	7	7	7	7
PD	29	29	29	29	28	27
DD	46	46	46	42	37	27
LD	180	180	178	103	52	9
LE	17	17	13	0	0	0

Table 2b Percentage of road segments at different scales in different types.

	1:1,000	1:5,000	1:20,000	1:50,000	1:100,000	1:200,000
EX	100%	100%	100%	100%	100%	100%
UT	100%	100%	100%	100%	100%	100%
PD	100%	100%	100%	100%	96.6%	93.1%
DD	100%	100%	100%	91.3%	80.4%	58.7%
LD	100%	100%	98.9%	57.2%	28.9%	5.0%
LE	100%	100%	76.5%	0%	0%	0%

Table 3 Mean road lengths at different portions of total number of roads at different scales.

	1:1,000	1:5,000	1:20,000	1:50,000	1:100,000	1:200,000
100%	3032.4	3032.4	3082.6	3615.9	4224.4	5134.3
90%	1271.6	1271.6	1329.4	1684.5	2109.2	2489.9
80%	665.8	665.8	690.2	1145.8	1540.5	1918.9
70%	435.1	435.1	457.2	734.7	1089.5	1510.6
60%	267.4	267.4	278.3	538.1	769.6	1174.1
50%	192.2	192.2	201.7	392.4	583.2	857.3
40%	130.9	130.9	134.8	261.7	453.2	603.7
30%	101.9	101.9	106.9	187.8	296.8	447.6
20%	66.9	66.9	70.4	124.6	170.5	205.1
10%	44.1	44.1	47.6	66.3	109.5	114.3

Table 4 Mean road widths at different portions of total number of roads at different scales.

	1:1,000	1:5,000	1:20,000	1:50,000	1:100,000	1:200,000
100%	26.4	26.4	26.8	30.7	33.6	38.3
90%	15.5	15.5	15.7	18.2	20.9	25.5
80%	11.8	11.8	12.1	14.3	15.8	20.6
70%	9.9	9.9	10.0	12.0	12.9	16.2
60%	8.2	8.2	8.3	10.1	10.6	13.5
50%	7.4	7.4	7.4	8.8	9.3	12.0
40%	6.7	6.7	6.8	7.6	7.7	10.4
30%	6.1	6.1	6.2	7.0	7.1	9.5
20%	4.8	4.8	5.0	5.4	5.0	7.7
10%	3.3	3.3	3.4	3.1	3.0	5.2

Table 5a Number of road segments at different scales for different lanes.

No. of Lanes	Number of road segments appears at different scales					
	1:1000	1:5000	1:20,000	1:50,000	1:100,000	1:200,000
1	86	86	85	28	19	5
2	106	106	105	88	58	30
3	23	23	23	21	14	9
4	25	25	25	24	15	8
5	8	8	8	8	7	7
6	7	7	7	7	7	7
7	1	1	1	1	1	1
8	2	2	2	2	2	2
9	1	1	1	1	1	1
10	0	0	0	0	0	0
11	1	1	1	1	1	1

Table 5b Number of road segments (percentage) at different scales for lanes from 1 to 5.

Number of lanes	Number of road segments (in percentage) appears at different scales					
	1:1000	1:5000	1:20,000	1:50,000	1:100,000	1:200,000
1	100.0%	100.0%	98.8%	32.6%	22.1%	5.8%
2	100.0%	100.0%	99.1%	83.0%	54.7%	28.3%
3	100.0%	100.0%	100.0%	91.3%	60.9%	39.1%
4	100.0%	100.0%	100.0%	96.0%	60.0%	32.0%
5	100.0%	100.0%	100.0%	100.0%	87.5%	87.5%

Table 6a Number of road segments at different scales for traffic ways from 0 to 3.

Number of Traffic Ways	1:1,000	1:5,000	1:20,000	1:50,000	1:100,000	1:200,000
0	20	20	16	1	0	0
1	111	111	111	62	35	21
2	147	147	145	117	88	48
3	2	2	2	2	2	2

Table 6b Number of road segments (in percentage) at different scales for traffic ways from 0 to 3.

Number of Traffic Ways	1:1000	1:5000	1:20,000	1:50,000	1:100,000	1:200,000
0	100.0%	100.0%	80.0%	5.0%	0.0%	0.0%
1	100.0%	100.0%	100.0%	55.9%	31.5%	18.9%
2	100.0%	100.0%	98.6%	79.6%	59.9%	32.7%
3	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 7a Number of road segments appeared in different scales for different orders.

Orders in segment	1:1000	1:5000	1:20,000	1:50,000	1:100,000	1:200,000
0,0	3	3	3	0	0	0
0,1	12	11	11	10	6	1
0,2	49	47	47	21	11	4
0,3	3	3	3	2	1	0
0,4	0	-	-	-	-	-
0,5	2	2	2	2	2	1
1,1	23	23	23	21	21	18
1,2	52	52	52	43	33	23
1,3	6	6	6	6	5	3
1,4	4	4	4	4	3	3
1,5	1	1	1	1	1	0
2,2	97	95	95	48	27	12
2,3	19	19	19	15	9	1
2,4	2	2	2	2	2	2
2,5	2	2	2	2	2	2
3,3	3	3	3	2	0	0
3,4	1	1	1	1	1	1
3,5	0	-	-	-	-	-
4,4	0	-	-	-	-	-
4,5	1	1	1	1	1	1
5,5	0	-	-	-	-	-

Table 7b Number of road segments in different scales categorized by their orders (after reclassified).

Orders in seg.	1:1,000	1:5000	1:20,000	1:50,000	1:100,000	1:200,000
0,0	3	3	3	0	0	0
X,0	66	66	63	35	20	6
1,1	23	23	23	21	21	18
X,1	63	63	63	54	42	29
X,X	125	125	123	71	42	18

Table 7c Number of road segments (in percentages) in different scales categorized by their orders (after reclassified).

Orders in seg.	1:1,000	1:5000	1:20,000	1:50,000	1:100,000	1:200,000
0,0	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%
X,0	100.0%	100.0%	95.5%	53.0%	30.3%	9.1%
1,1	100.0%	100.0%	100.0%	91.3%	91.3%	78.3%
X,1	100.0%	100.0%	100.0%	85.7%	66.7%	46.0%
X,X	100.0%	100.0%	98.4%	56.8%	33.6%	14.4%

Table 8 Highest degree (academic qualification) received and perception of life.

			LIFE			Total
			Dull	Routine	Exciting	
DEGREE (qualification)	High School	Count Row %	35 6.8%	251 48.5%	231 44.7%	517 100%
	Bachelor	Count Row %	2 1.3%	58 36.9%	97 61.8%	157 100%
	Graduate	Count Row %	1 1.4%	21 28.8%	51 69.9%	73 100%
Total		Count Row %	38 5.1%	330 44.2%	379 50.8%	747 100%

Table 9 Values of variable life and degree.

	Life	Degree
Case 1	1	2
Case 2	2	3
Case 3	3	2

Table 10 Number of road segments at different FVL in different portions of length.

Road length (l) in kilometers	Feature Vanishing Levels (FVL)						Total
	1	2	3	4	5	6	
$0 < l \leq 0.1$	0	5	40	14	2	1	62
$0.1 < l \leq 0.2$	0	0	36	11	8	10	65
$0.2 < l \leq 0.3$	0	1	11	12	9	2	35
$0.3 < l \leq 0.4$	0	0	2	6	3	3	14
$0.4 < l \leq 0.5$	0	0	3	3	6	3	15
$0.5 < l \leq 0.6$	0	0	1	3	7	4	15
$0.6 < l \leq 0.7$	0	0	0	2	3	5	10
$0.7 < l \leq 0.8$	0	0	0	0	2	0	2
$0.8 < l \leq 0.9$	0	0	0	0	3	5	8
$0.9 < l \leq 1.0$	0	0	0	1	1	4	6
$1.0 < l \leq 1.1$	0	0	0	0	1	0	1
$1.2 < l \leq 1.3$	0	0	0	0	1	4	5
$1.3 < l \leq 1.4$	0	0	0	2	0	2	4
$1.4 < l \leq 1.5$	0	0	0	0	1	2	3
$1.5 < l \leq 1.6$	0	0	0	0	0	2	2
$1.6 < l \leq 1.7$	0	0	0	2	2	3	7
$1.7 < l \leq 1.8$	0	0	0	0	1	1	2
$1.8 < l \leq 1.9$	0	0	0	0	1	2	3
$1.9 < l \leq 2.0$	0	0	0	0	1	1	2
$2.2 < l \leq 2.3$	0	0	0	0	0	3	3
$2.3 < l \leq 2.4$	0	0	0	0	1	1	2
$2.4 < l \leq 2.5$	0	0	0	0	0	2	2
$2.6 < l \leq 2.7$	0	0	0	0	0	3	3
$3.4 < l \leq 3.5$	0	0	0	0	1	0	1
$3.5 < l \leq 3.6$	0	0	0	0	0	2	2
$3.6 < l \leq 3.7$	0	0	0	0	0	1	1
$3.7 < l \leq 3.8$	0	0	0	0	0	1	1
$4.3 < l \leq 4.4$	0	0	0	0	0	1	1
$6.9 < l \leq 7.0$	0	0	0	0	0	2	2
$8.4 < l \leq 8.5$	0	0	0	0	0	1	1
Total	0	6	93	56	54	71	280

Table 11 Number of road segments at different FVL in different portions of width.

Road width (*w*) In meters	Feature Vanishing Levels (FVL)						Total
	1	2	3	4	5	6	
$0< w \leq 5$	0	5	17	7	16	3	48
$5< w \leq 10$	0	1	64	28	26	20	139
$10< w \leq 15$	0	0	10	13	7	21	51
$15< w \leq 20$	0	0	2	8	5	8	23
$20< w \leq 25$	0	0	0	0	0	8	8
$25< w \leq 30$	0	0	0	0	0	4	4
$30< w \leq 35$	0	0	0	0	0	2	2
$35< w \leq 40$	0	0	0	0	0	3	3
$45< w \leq 50$	0	0	0	0	0	1	1
$50< w \leq 55$	0	0	0	0	0	1	1
Total	0	6	93	56	54	71	280

Table 12 Number of road segments at different FVL for different connectivity of road.

Connectivity	Feature Vanishing Levels (FVL)						Total
	1	2	3	4	5	6	
0,0	0	0	2	0	0	0	2
X,0	0	4	28	15	14	6	67
1,1	0	0	2	0	3	18	23
X,1	0	0	9	12	13	29	63
X,X	0	2	52	29	24	18	125
Total	0	6	93	56	54	71	280

Table 13 Number of road segments at different FVL for different types of road.

Type	Feature Vanishing Levels (FVL)						Total
	1	2	3	4	5	6	
Non-transport use	0	0	2	0	0	0	2
Lane	0	4	9	0	0	0	13
Local Distributor	0	2	77	51	41	7	178
District Distributor	0	0	5	4	10	29	48
Primary Distributor	0	0	0	1	3	27	31
Trunk Road	0	0	0	0	0	8	8
Total	0	6	93	56	54	71	280

Table 14 Number of road segments at different FVL for different number of traffic way.

Number of Traffic way	Feature Vanishing Levels (FVL)						Total
	1	2	3	4	5	6	
0	0	4	15	1	0	0	20
1	0	0	49	27	14	21	111
2	0	2	29	28	40	48	147
3	0	0	0	0	0	2	2
Total	0	6	93	56	54	71	280

Table 15 Number of road segments at different FVL for different number of lane.

Number of lane	Feature Vanishing Levels (FVL)						Total
	1	2	3	4	5	6	
1	0	1	57	9	14	5	86
2	0	1	17	30	28	30	106
3	0	0	2	7	5	9	23
4	0	0	1	9	7	8	25
5	0	0	1	0	0	7	8
6	0	0	0	0	0	7	7
7	0	0	0	0	0	1	1
8	0	0	0	0	0	2	2
9	0	0	0	0	0	1	1
11	0	0	0	0	0	1	1
Total	0	2	78	55	54	71	260

Table 16 Dependence of FVL on road attributes by Somers' Δ and Lamda tests

Tests	Value	Approximate Significance
FVL Dependency on road length	Δ = 0.529	0.0005
FVL Dependency on road width	Δ = 0.357	0.0005
FVL Dependency on road type	Δ = 0.733	0.0005
FVL Dependency on number of traffic ways	Δ = 0.409	0.0005
FVL Dependency on number of lanes	Δ = 0.481	0.0005
FVL Dependency on connectivity	λ = 0.193	0.0005