# Study on the Method of Complex Lines Selection and Automated Processing for Railway Marshalling Yard

Chengming Li, Xixia You, Xiaoli Liu and Pengda Wu
**Chinese Academy of Surveying and Mapping, Beijing, China**

*Abstract*—**As the central of railway hub, railway marshalling yard is the base for vehicle distribution and disintegration. Time-consuming of operation and residence for vehicle is common to spend 70% of the goods turnover cycle, thus the railway line here is extremely complex. Meanwhile, there are few articles talking about the generalization selection method for marshalling yard because of the small number and a lot of existing research results on road are not applicable to it. So, this paper presents a new automation generalization method for railway marshalling yard. Firstly, six map features are extracted based on its spatial distribution and topological characteristics, including broken line, the same node line, similar line, mesh line, parallel line and main line. Secondly, three basic principles suitable for generalization selection and automation processing method are put forward. Finally, practical data is used to validate the method and the results show that it is reasonable and effective.**

*Keywords-generalization selection; railway marshalling yard; map features; basic principles for generalization selection*

## I. INTRODUCTION

The marshalling yard is a type of traffic component. Early content selection only considers the level and attribute of railway lines, namely, standard monorail (double track), narrow-gauge railway (double track), and narrow-gauge railway (single track). In case of special requirements, content selection is performed by following the condition-conclusion rule. Marc van Kreveld et. al. propose that on the basis of level-based traffic route selection, computation should be used to determine whether the distance between the reserved route node and accidental node is smaller than tolerance and that new routes should be added to avoid detours so long as no new conflicts arise. For the map data of most marshalling yards, the internal lines are not assigned with any attribute values. Obviously, attribute-based selection methods are hardly applicable to marshalling yards. In addition, common traffic route selection methods include those based on intensity values of mesh nodes and road generalization methods based on genetic algorithm. However, apart from line networks in marshalling yards, there are complex map features including suspension lines and parallel arcs. Practice has proved that the performance of these line selection methods is not ideal.

As there are only less than a hundred marshalling yards in China, researchers have not given sufficient attention to this problem. Meanwhile, most research results concerning traffic routes and river systems are not applicable to such elements. During automatic map generalization, it is impossible to skip these elements like marshalling yards. Moreover, selection methods designed for other linear objects tend to perform poorly on marshalling yards. Therefore, based on the analysis of the spatial layout and topological features of marshalling yard, this paper summarizes six map features including broken arc, same-node arc, similar arc, mesh arc, parallel arc and main arc, and proposes three major principles for the generalization selection of marshalling yards. With these findings, an automatic processing method is devised. Finally, the rationality of these principles and the effectiveness of selection methods are validated through practical data tests.

## II. MAP FEATURES OF MARSHALLING YARD

The clearly defined use of the marshalling yard determines its spatial layout and structure. According to the difference between shunting systems and between arrangements of departure yards, receiving yards and shunting yards, marshalling yards can be classified into three spatial structures: longitudinal-type, transverse-type and hybrid-type. All other structures can be seen as derivatives of the three structures above.



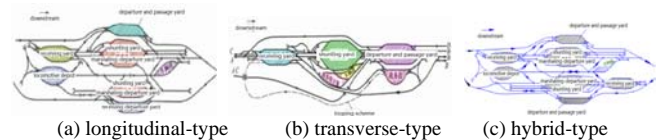(a) longitudinal-type      (b) transverse-type      (c) hybrid-type

FIGURE I. THREE STRUCTURAL FEATURES OF MARSHALLING YARD

For the ease of subsequent automatic content selection, this work summarizes six map features from lines in marshalling yards: broken arc, same-node arc, similar arc, mesh arc, parallel arc and main arc. According to the topological features in Figure II, a topological relationship tree and a topological information table concerning points, lines and planes are established. The topological relationship tree and the topological information table can clearly demonstrate the topological relations and connectivity between nodes and arcs.

Broken arc. In a marshalling yard, broken arcs are arcs whose head and end nodes correlate with no arcs, namely, independent arcs in topological features. These arcs are mainly used to store obsolete wagons and equipment that can be used as replacement parts. The connectivity of their head and end nodes is "1", without any left or right polygons. L18 and L19 in Figure II are broken arcs.
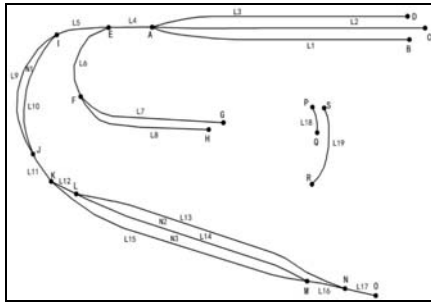
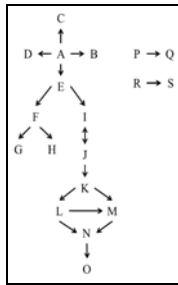FIGURE II. SKETCH OF SPATIAL CHARACTERISTICS OF MARSHALLING YARD'S INTERNAL



FIGURE III. TOPOLOGICAL RELATIONSHIP TREE OF NODES

Same-node arc. In a marshalling yard, same-node arcs are arcs with the same head and end nodes with a small included angle. Same-node arcs are mainly used for transferring cargo between two wagons. The connectivity of their nodes is "3" with left and right polygons. One side of same-node arcs is empty while the other side is not. L9 and L10 in Figure II are same-node arcs.

Similar arc. In a marshalling yard, similar arcs are two suspended arcs that are connected by one node and have a small included angle and distance between them. Similar arcs are mainly used for transferring cargo between multiple wagons. For similar arcs, the connectivity of one node is "3" while the connectivity of the other node is "1". The left and right polygons are both empty. L7 and L8 in Figure II are similar arcs.

Mesh arc. In a marshalling yard, mesh arcs are middle arcs that form a closed polygon. Lines in receiving yards, shunting yards and departure yards are often mesh arcs. Mesh arcs are formed by one or more adjacent same-node arcs, such as the arcs that form N1, N2 and N3 polygons in Figure II.

Parallel arc. Among the supporting facilities in a marshalling yard, many lines are close and almost parallel to each other. Parallel arcs are often made up by multiple suspension arcs. These arcs are mainly used for parking wagons and performing operations. For parallel arcs, the connectivity of one node is "3" while the connectivity of the other node is "1". The left and right polygons are both empty. L1, L2 and L3 in Figure II are parallel arcs.

Main arc. In a marshalling yard, main arcs are arcs that are connected to the receiving yard, shunting yard and departure yard. They are often made up by non-mesh middle arcs and mainly used for wagon arrival and departure. For main arcs,

the connectivity of their both ends is greater than or equal to "2". The left and right polygons are both empty. L5 and L11 in Figure II are main arcs.

## III. SELECTION PRINCIPLES AND AUTOMATIC PROCESSING METHODS

When the map data of a marshalling yard is generalized from an original scale 1:Scale1 to a target scale 1:Scale2 (Scale2>Scale1), which lines should be selected and which should be discarded are naturally correlated with the original scale and target scale as well as their own structural features and topological features in the marshalling yard. In other words, the length threshold and distance threshold for generalization selection are determined by the original scale and the target scale. When several arcs all meet the selection conditions, the contribution and value of their structural robustness and topological extensibility become deciding factors. Therefore, generalization selection should follow or meet some basic principles and requirements such as visual conspicuity, structural robustness and topological extensibility.

### A. Visual Conspicuity

The marshalling yard is composed of departure yards, shunting yards, receiving yards and other components which are also the areas to be processed during generalization selection. The analysis of large amounts of marshalling yard data shows that the effective length of normal departure line, shunting line and receiving line, depending on vehicle capacity, ranges between 600 to 1,000 meters. The other lines (except main lines) are mostly shorter than 600 meters. Therefore, the generalization selection for marshalling yards often involves 1:500~1:25000 scale. Scales smaller than 1:25000 can only reserve the rough shapes of the marshalling lines, obviating the need to show internal lines. In the existing studies, length threshold and distance threshold for the generalization from the original scale to the target scale are empirical values, which are unsuitable for automatic processing. According to domestic and foreign cartographic representation standards, on a 1:25000-scale map, tunnels, mingtong tunnels, dams, steps, and land boundaries should be no shorter than 1 mm while station-line distance should be no shorter than 0.3 mm. Otherwise, it would be difficult for human eyes to distinguish them. From the perspective of information measurement, reserving these components will greatly increase the complexity and produce redundant information. Based on this analysis, the model for determining length threshold for 1:Scale (Scale>25000) scale can be deducted (see Equation 1 and 2) while the model for determining distance threshold can be deducted (see Equation 3 and 4).

### 1) Model for determining the length threshold

$$T_{1:Scale} = \frac{1:Scale}{1:25000} \times T_{1:25000}, \quad T_{1:25000} = 1mm \tag{1}$$

where $T_{1:25000}$ is the smallest distinguishable length on a 1:25000-scale map and is often assigned 1mm as its value; 1:Scale is the scale of a certain map; $T_{1:scale}$ is the smallest distinguishable length on a 1:Scale map.

$$LT_{1:Scale1 \rightarrow 1:Scale2} = \left(1 - \frac{1:Scale2}{1:Scale1}\right) \times T_{1:Scale2} \tag{2}$$

Where $LT_{1:scale1 \rightarrow 1:scale2}$ is the length threshold for generalization selection on a 1:Scale1 map; 1:Scale1 is the original scale of map generalization selection; 1:Scale2 is the target scale of map generalization selection; $T_{1:scale2}$ is the smallest distinguishable length on a 1:Scale2 map and is obtained by computing Equation 1.

*2)    Model for determining the distance threshold*

$$SpaceT_{1:scale} = \frac{1:Scale}{1:25000} \times SpaceT_{1:25000}, \quad SpaceT_{1:25000} = 0.3mm \tag{3}$$

where $SpaceT_{1:25000}$ is the smallest distinguishable distance on a 1:25000-scale map and is often assigned 0.3 mm as its value; 1:Scale is the scale of a certain map; $SpaceT_{1:scale}$ is the smallest distinguishable distance on a 1:Scale map.

$$Th_{1:Scale1 \rightarrow 1:Scale2} = \left(1 - \frac{1:Scale2}{1:Scale1}\right) \times SpaceT_{1:Scale2} \tag{4}$$

where $Th_{1:scale1 \rightarrow 1:scale2}$ is the distance threshold for generalization selection on a 1:Scale1 map; 1:Scale1 is the original scale of map generalization selection; 1:Scale2 is the target scale of map generalization selection; $SpaceT_{1:scale2}$ is the smallest distinguishable distance on a 1:Scale2 map and is obtained by computing Equation 3.

*3)    Calculation examples*

The original scale of the data in Figure II is 1:500 while the target scale is 1:10000. By calculating Equation (1), we get the smallest distinguishable length $T_{1:10000}$=2.5mm on a 1:10000 map. By calculating Equation (2), we get the length threshold $LT_{1:500 \rightarrow 1:10000}$ = 2.375mm for generalization selection. By calculating Equation (3), we get the width threshold $SpaceT_{1:10000}$= 0.75mm on the map. By calculating Equation (4), we get the distance threshold $Th_{1:500 \rightarrow 1:10000}$= 0.7125mm for generalization selection.

## B.    Structural Robustness

After length threshold and distance threshold are obtained, how is the selection performed when the mesh arcs and similar arcs are both smaller than the thresholds? It is crucial to maintain the original structure to the greatest extent possible. During manual implementation, intuitive qualitative analysis should be used to reserve important, conspicuous backbone lines and delete unimportant, discrete lines. Computer-aided automatic processing mainly rely on thresholds and quantifiable structural features and topological features.

*1)    Processing of broken arcs*

According to the geometric length of broken arcs, arcs shorter than length threshold are put into the container. After all broken arcs are processed, the arcs in the container are deleted from the topological structure. Then, the topological structure is updated to obtain new topological relationships.

For example, the broken arc L18 is deleted because its geometric length D18 = 13m/10000=1.3mm is smaller than the length threshold 2.375mm. But L19 is reserved. Discrete lines are deleted to ensure the performance of generalization selection for overall structure.

*2)    Processing of parallel arcs*

According to Table I (a) and the definition of map features, we can find suspension arcs from the topological arcs. Sampling points that can represent basic arc features are selected to form a new element Arc. A buffer zone is created for the arcs using distance threshold as the buffer radius. Then, by querying all arcs intersecting the buffer zone, we get Arc1 and Arc2, two arcs nearest to Arc on its two sides as well as the distance between Arc1 and Arc2. If the distance is smaller than the threshold, then the current arc should be deleted and the topological relationships should be updated. The distance $Space_{arc1 \rightarrow arc2}$ between Arc1 and Arc2 can be calculated using Equation (5).

$$Space_{arc1 \rightarrow arc2} = \frac{\sum_{i=1}^{N_1} D_i + \sum_{i=1}^{N_2} L_i}{N_1 + N_2} \tag{5}$$

where N1 and N2 are the numbers of sampling points on the two arcs, respectively. Di and Li are the distances from a sampling point on one arch to the other arch.

After the buffer zone for L2 is created based on distance threshold, we can obtain two arcs in the zone, L1 and L3. By calculating the distance between L1 and L3, we get S=14m/10000=1.4mm which is greater than the distance threshold 0.7125mm; the distance between L1 and L2, L2 and L3 is S=7m/10000=0.7mm, which is smaller than the distance threshold 0.7125mm. So, L2 should be deleted while L1 and L3 should be reserved to ensure than the main structure is unchanged.

*3)    Processing of same-node arcs and similar arcs*

These two map features are both characterized by a small included angle and similar shapes. Consequently, deleting any one arch would not affect the main structure of the marshalling yard. During generalization, one reserved arch is sufficient to represent the ground object feature. Through the calculation of Equation (5), we can determine the distance between the two. If the distance is smaller than the threshold, the selection should be performed based on the geometric length of the two arcs. In that case, the longer arch should be reserved while the shorter one should be deleted. Meanwhile, the topological relationships should be timely updated.

For the same-node arcs L9 and L10 in Figure II, as L9 is longer than L10, L9 is reserved and L10 is deleted. Similarly, for the similar arcs L7 and L8, as L7 is longer than L8, L7 is reserved and L8 is deleted.

## C.    Topological Extensibility

During generalization selection, the topological extensibility of the marshalling yard should be guaranteed. Although some arcs are short and narrow, we still need to reserve arcs that are important for maintaining the connectivity

in the marshalling yard and discard lines that have less connectivity and correlation with other lines. Naturally, the positional relations between line should be clearly and rationally demonstrated. Following this principle, we can perform automatic processing for main lines and mesh arcs.

### 1) Processing of main arcs

By definition, main arcs are arcs whose connectivity of head and end nodes is greater than or equal to "2" and whose left and right polygons are both empty. With this definition, we can determine main arcs in a marshalling yard. As main arcs serve the role of connecting receiving yards, shunting yards and departure yards, they are made up by non-mesh middle arcs. Therefore, main arcs play a crucial role in maintaining topological connectivity and thus should be reserved.

### 2) Processing of mesh arcs

There are often shared arcs between two adjacent meshes. But deleting shared arcs would not affect the connectivity of the remaining nodes in the mesh. So, the major task for the processing of mesh arcs is to delete shared arcs. According to the plane's topological structure, all net polygons are put into a container for processing in turn. For single mesh, if it is formed by two arcs, then the processing is the same as that for same-node arcs. If it is made up by more than two arcs, then the arcs are processed starting from the longest one to the smallest one.

### D. Workflow of Automatic Selection

The workflow of automatic generalization selection for the marshalling yard can be divided into three procedures. First, length threshold and width threshold are determined according to the original scale and target scale of the marshalling yard. Second, the topological relationships of marshalling yard are constructed to obtain the topological correlation and connectivity information of nodes, lines and planes. Third, the six map feature arcs in the marshalling yard are processed by following the principles mentioned above. The processing order should be: short broken arcs should be first processed, then mesh arcs and parallel arcs, and finally the remaining same-node arcs and similar arcs. During the construction of topological relationships, basic data should be topologically pre-processed, including the processing of arc self-intersection, node fitting, omission of repetitive lines, false nodes, redundant nodes, and short suspended lines, and the construction of topological polygons. The pre-processed data is used to create a topological correlation table to illustrate the topological relationship between network structure elements (nodes, arcs and planes) in the form of a table (see Table I). Consequently, spatial features such as broken arcs, same-node arcs, similar arcs and mesh arcs can be identified. Figure IV shows the selection results of automatic generalization from 1:500 scale to 1:10000 scale of Figure II.
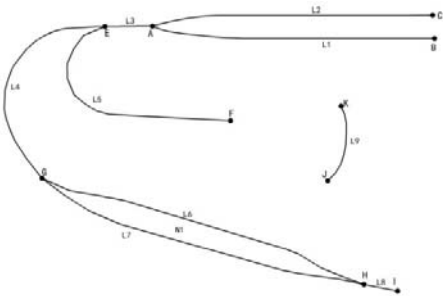


FIGURE IV. SELECTION RESULTS OF AUTOMATIC GENERALIZATION FROM 1:500 SCALE TO 1:10000 SCALE OF FIGURE II

## IV. EXPERIMENT AND ANALYSIS

This paper takes Chengdu East Railway Station, a 30 sq km hybrid-type marshalling yard, as the object of the research. The map data is manually collected from aerial photos taken in 2015. The original scale is 1:500 with a total of 611 arcs. The original image is shown in Figure V (a). This marshalling yard is composed of receiving yard, shunting yard, departure yard, and other components, and includes all six map features. The experiment is conducted on the WJ-III map workstation developed by Chinese Academy of Surveying and Mapping and embedded with the proposed generalization selection method, where the original 1:500 topographical map of the marshalling yard is generalized into a 1:10000 scale map. The key parameters in automatic generalization selection are shown in Table I. The results of generalization are shown in Figure V (b).
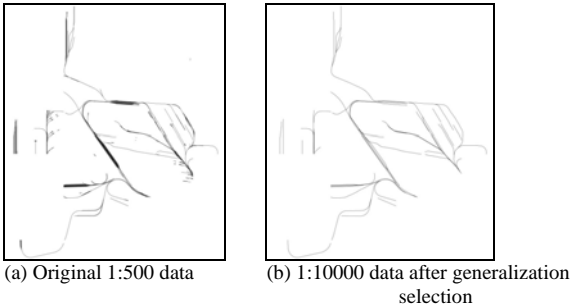


(a) Original 1:500 data      (b) 1:10000 data after generalization selection

FIGURE V. COMPARISON OF THE RESULTS BETWEEN BEFORE AND AFTER

TABLE I. KEY PARAMETERS IN AUTOMATIC GENERALIZATION SELECTION

| Scale | Number of arcs | Length threshold | Distance threshold | Processing time |
|---|---|---|---|---|
| Before (1:500) | 611 | 2.375mm | 0.7125mm | 0.256s |
| After (1:10000) | 99 | | | |

By analyzing the parameters above and comparing the two topographical maps, we can easily make the following findings. First the whole automatic processing time is 0.256s, which is highly efficient. Second, the selection performance is outstanding and reserves only 99 arcs (16.2% of the original arcs). Third, the main structure of the marshalling yard is maintained. The important connectivity between stations is left

intact. Dense arcs are simplified and discrete and scattered arcs are deleted.

## V. CONCLUSION

Based on the analysis of the marshalling yard's spatial layout and topological features, this paper summarizes six map features, including broken arc, same-node arc, similar arc, mesh arc, parallel arc and main arc, and proposes three major principles for marshalling yard generalization selection. The formula for calculating length threshold and distance threshold are deduced. Meanwhile, this study designs an automatic selection method for the generalization of the six map features and embeds the propose selection algorithm into the WJ-III map workstation developed by Chinese Academy of Surveying and Mapping. The 1:500 map data of the 30-sq-km hybrid-type Chengdu East Railway Station is used for the experiment of automatic generalization selection. The results show that the reserved arcs accounts for only 16.2% of the original arcs and that the main structure of the marshalling yard is complete and has good topological connectivity.

However, this paper only realizes the automatic generalization selection for marshalling yard but does not look at arc simplification and shift or integrate with the attribute data of marshalling yard. These issues still awaits further discussion.

## REFERENCES

[1] Wang Guangxia. Research of Auto-generalization System and Realization of Auto-generalization for Habitation and Road [D]. Zhengzhou: Institute of Surveying and Mapping of PLA, 1994.

[2] van Kreveld M, Luo J. The definition and computation of trajectory and subtrajectory similarity[C]//Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems. ACM, 2007: 44.

[3] WANG Jiayao, CUI Tiejun, WANG Guangxia. Applications of Graph Theory in Automatic Selection of Road Network [J]. Journal of Geomatics Science and Technology, 1985 (1): 79-86.

[4] HU Yungang, CHEN Jun, LI Zhilin, ZHAO Renliang. Selective Omission of Road Features Based on Mesh Density for Digital Map Generalization[J]. ActaGeodaetica et Cartographic Sinica, 2007, 36(3): 351.

[5] TIAN Jing, SONG Zihang, AI Tinghua. Grid Pattern Extraction in Road Networks with Graph[J]. Geomatics and Information Science of Wuhan University, 2012, 37 (6): 724-727.

[6] DENG Hongyan, WU Fang, ZHAI Renjian, LIU Weiwei. A General ization Model of Road Networks Based on Genetic Algorithm[J]. Geomatics and Information Science of Wuhan University, 2006, 31(2): 164-167.

[7] ZHAI Renjian, WU Fang, DENG Hongyan, et. al. An Automated Selection Model of Ditch Based on Multi-objective Optimization by GeneicAlgorithm[J]. ActaGeodaetica et Cartographic Sinica, 2008, 37(1): 108-331.

[8] WANG Yanhui, CHEN Jun. Interpolation Method of the Road Object Selection[J]. ActaGeodaetica et Cartographic Sinica, 2008, 37(2): 78-83.

[9] Benz S A, Weibel R. Road network selection for medium scales using an extended stroke-mesh combination algorithm[J]. Cartography and Geographic Information Science, 2014, 41(4): 323-339.

[10] WANG Pengbo, WU Fang, ZHAI Renjian. A Topologic Operation Method for Automated Generalization of Road Networks[J]. Journal of Geomatics Science and Technology, 2009, 26(1): 64-68.

[11] CHEN Bo, WU Fang, QIAN Haizhong. Spatial Analysis in Automatic Topological Construction for Road Network[J]. ActaGeodaetica et Cartographic Sinica, 2006, 31(1): 101-102.

[12] LI Fei, LUAN Xuechen, YANG Bisheng, LI Qiuping. Automatic Topology Maintenance Approach for High-level Road Networks[J]. Geomatics and Information Science of Wuhan University, 2014, 39(6): 729-733.

[13] YANG Bisheng, LUAN Xuechen, LI Qiuping. Generating Hierarchical Strokes from Urban Street Networks Based on Spatial Pattern Recognition[J]. International Joural of Geographical Information Science, 2011,25(12):2025-2050.

[14] YU Z. The Effects of Scale Change on Map Structure[D].Worcester: Clark University,1993,27-61.

[15] HE Jianhua, LIU Yaolin. An Integrated Model for Topology &Direction Relation Reasoning[J]. ActaGeodaetica et Cartographic Sinica , 2004, 33(2): 156-162.