

# Matching Road Network Combining Hierarchical Strokes and Probabilistic Relaxation Method

Lin Yang, Zejun Zuo\*, Run Wang, Yaqin Ye and Maosheng Hu

Faculty of Information Engineering, China University of Geosciences (Wuhan), Wuhan, 430074, China, State Key Laboratory of Geography Information Engineering, Xi'an 710054, China

**Abstract:** Aiming at the complex multisource road network matching modes, this paper proposed a road network matching method based on hierarchical stable strokes. The progress of the matching method adopts recursive method and gives consideration to both global consistency and local similarity of homonymous road features. One layer of hierarchical strokes generating from road spatial structure are selected at a match and the matched parts are used as a stable reference for the next layer. The detailed matching is implemented by an iterative probability relaxation method. In the course of coarse matching three similarity indicators, including the geometry shape similarity of strokes, the relative position of nodes and strokes and the topological structure similarity of nodes, are introduced into the initial matching probability calculation. After that, an iterative calculation is carried out for updating matching probability. Finally, a self-designed selecting principle is applied to trace the exact matching results. Matching road centerline from different producers of the same geographical area shows that our method finds 1:1, 1:N and M:N matching modes and gets a satisfactory matching result, which means hierarchical stable reference and the stroke structure can provide an effective way to identify different matching relations and avoid the non-rigid deviation of the homonymous features.

**Keywords:** Hierarchical strokes, probability relaxation method, road network matching, stable strokes.

## 1. INTRODUCTION

Road network spatial target matching is a core and key technology of application that vector road map such as integration, update and navigation based on location service. Specifically, it is the process of identifying the same object or feature in the real world expressed in the different sources of map data through the analysis of a series of spatial entity similarity indexes [1]. Road network matching technology can not only serve the application of the vector road map integration, reduce the high cost of the repetitive data acquisition, improve the quality of spatial data; but also satisfy the strong demand of the rapid development of vehicle navigation products for the present situation of navigation electronic map data. It is a hot problem both at home and abroad.

Different scale, time and acquisition processes lead to the data difference on scale, precision and data content. Therefore, both road segments and intersection nodes are likely to have multiple types of correspondence which presents a complicated match relationship. For this reason, road network matching is actually the simulation of very complex human thinking activity.

In view of the road network matching problem, many scholars have adopted road elements matching method based on geometric features that uses the indexes of distance, length, direction, shape, and so on, for geometric similarity evaluation and measure. In light of the natural topological

characteristics of road network, comprehensive measurement method with the combination of the geometry and topology is developing [2-12]. Matching methods faced some common problems for a long time: the excessive attention to local shape of spatial target. Over-reliance on local morphological similarity makes matching easy to fall into local optimization, which makes poor performance in the situation of larger non-systematic deviation, and it is difficult to handle the matching relations of  $1:N$  or  $M:N$  types, then matching based on global consistency began to appear in the road network matching [13-18]. DB Zhao [13] starts from the overall consistency when considering the similarity index of various features to match the situation, taking into account the elements of the neighborhood, to obtain a more precise matching relationship, the method focuses on  $1:N$  correspondence. YF Zhang and BS Yang [14, 15] proposed a probabilistic matching unit relaxation method to solve the  $1:0$  matching problem and two-way matching problem. Zhang [16] introduced the concept of stroke to matching, improved mismatch situation existing in local matching method. YG Hu, J Chen, *et al.* [17] proposed a hierarchical organization divided by "basic, decomposition and abstraction" for road linear target.

Aiming at the local failure condition of the individual geometric similarity matching method, this paper will mine stable calculation space from the city road spatial structure to get more reference information, as the auxiliary similarity measure. In this paper, an urban road network matching method is proposed based on hierarchical stable strokes. In addition, the algorithm introduces the method of probabilistic

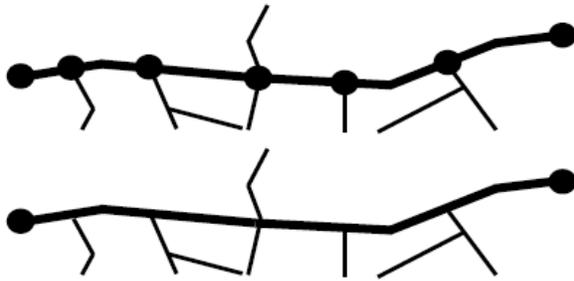


Fig. (1). Road segment and strokes Structure stroke structure.

relaxation to matching algorithm. According to the criterion of local pair point, each pair point on value of the similarity measure is adjusted parallel, using context information to reduce local fuzziness and achieving global consistency through iterative process. Finally, the actual road network data of Wuhan city is selected and the algorithm is verified.

2. THEORETICAL BASIS AND ALGORITHM

2.1. Hierarchical Stable Strokes

Algorithm uses stroke as the basic unit of matching. Stroke is a set of direction consistent and spatially continuous sections. As shown in Fig. (1), the six sections in (a) represented as a continuous strokes in (b). Using this structure on the one hand can keep roads connected and continuous, and helpful to find stable structure in space; on the other hand, taking strokes as matching unit, which can be easier to constrain the internal node matching from the whole aspect.

Urban road network has obvious hierarchies. Matching method based on the stable level strokes is dependent on this hierarchy of road network. Hierarchical method for processing is consistent with the road on the stability of high and low order, and in conformity with the road to the difficulty of the match. The stable reference space provides (1) Public reference: It is easier to calculate the transformation and the establishment of space. (2) Context information: Stroke information provides more context information for the matching relationship between the local elements. When the non-systematic deviation and inhomogeneous deformation is obvious, the structure information can be used to assist in determining the matching relations.

2.2. Algorithm’s Basic Idea

The progress of the matching method adopts recursive method and gives consideration to both global consistency and local similarity of homonymous road features as shown in Fig. (2). One layer in the hierarchy of the road data is selected at a match and the matched are used as a stable reference for the next layer. Our method starts with high grade roads and extracts matched stroke pairs of them as initial common “coordinate axis” (also known as common reference set or initial stable set). Because high grade roads can describe the structure of road network with only a few data and play a key role in connecting road segments. They have a virtue of strong stability for there is a small variation to

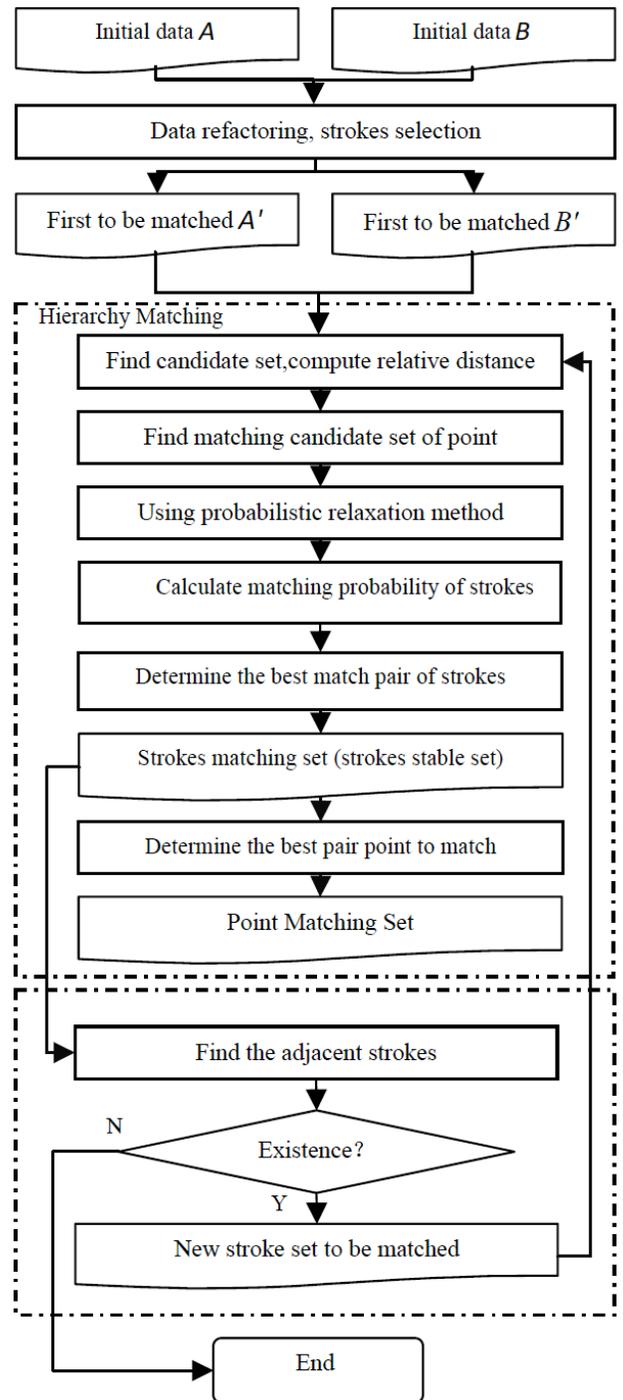


Fig. (2). Flow chart of the hierarchical matching method.

them in few years. Moreover, the space among high grade roads is usually large enough to reduce the possibility of the mismatch. After that, two new matched sets are generated by adjacent strokes of the two initial stable set. Through computing relative distance of the two sets, different source data are converted into the same computation space so that the influence of non-rigid deviation is lessened. Matching method also uses the shape similarity of stroke and the topological similarity of its internal nodes as indicators. The best matching sets are generated by an iterative probability relaxation

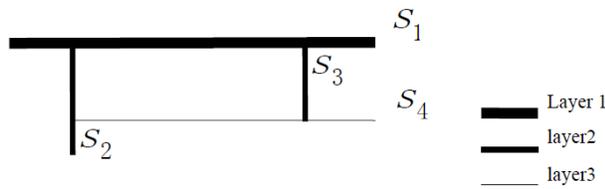


Fig. (3). Schematic diagram of road network hierarchy.

method, which is applied to determine the matching relation of strokes. The matching relations of nodes as  $1 : 1, 1 : N$  and  $M : N$  pairs are already determined when calculating the topological similarity of them. Finally, the best matched are added into the initial stable set to form a new stable set. A further search of adjacent strokes begins a new round of road matching. Repeating these steps until no more strokes can be selected for matching and then, the final stable set is the matching results. Matching road centerline from different producers of the same geographical area shows that our method is able to avoid the non-rigid deviation of the homonymous features, match  $1 : 1, 1 : N$  and  $M : N$  pairs correctly and get a stable matching result.

### 3. MATCHING METHOD BASED ON HIERARCHICAL STABLE STROKES

#### 3.1. The Hierarchical Strokes

The first layer strokes selects high-grade road through the data reconstruction in the road network. High-grade roads summarize the main structure of the road network by the smaller amount of data and play a supporting role in the skeleton of the road network with the smaller possibility of change and strong stability. Geometry distance between High-grade roads is large, and the possibility of mismatching is relatively low and better matches are obtained.

As shown in Fig. (3), stroke  $S_1$  is in the first layer. And adjacent strokes  $S_2$  and  $S_3$  are in the second layer of the matched sets. Similarly, the adjacent of  $S_2$  and  $S_3$  is  $S_4$ , which is the third layer to be matched. If a stroke does not find the matching pair within the layer to which they belong, the stroke will be downgraded to the next level.

#### 3.2. Matching within Layer Based on The Probabilistic Relaxation Method

Assume extracting two high-grade strokes sets: strokes set  $A'$  and strokes set  $B'$  from two road network data to be matched from A and B.  $Sa' = \{Pa_1, Pa_2, L, Pa_m\}$  are strokes in  $A'$ , and  $Pa_i$  is the internal topology point(endpoint and cross point). Firstly, analyze and determine the candidate matching set for  $Sa'$  in  $B'$  using a buffer analysis, denoted as  $B_{CS} \subseteq B'$ . If the situation  $B_{CS} = j$  exists, that is,  $Sa'$  does not have candidate matching set, so  $Sa'$  will be excluded from the original stable reference. Suppose  $Sb' = \{Pb_1, Pb_2, L, Pbn\} \in B_{CS}$  is a set of candidate roads then analyze the set of candidates

that matches with each point  $Pa_i$  in  $B_{CS}$  for strokes  $Sa'$  using buffer analysis, and then further matches every point in the first layer strokes using probabilistic relaxation method. The matching steps are as following:

1) Determine the initial probability.

First, set the initial matching probability between nodes within a layer. In addition to topological structure of point for the similarity between points, consider joining the strokes similarity constraints. Initial probability is measured by two indicators: the strokes similarity which point belongs  $Sim_{stroke}$  and the topological similarity of point  $Sim_{node}$ .

Suppose  $Pb_j \in B_{CS}$  is a candidate point of  $Pa_i$ , their similarity as shown in Equation (1).

$$Sim_{Pa_i, Pb_j} = e_1 Sim_{Stroke} + e_2 Sim_{node} \quad (1)$$

$Sim_{stroke}$  is the overall similarity of the strokes, using linear similarity  $Sim_{line}$  of matching point where the strokes to measure. In reference [13] linear similarity calculation method, suppose  $Sa'$  and  $Sb'$  are two strokes to be matched having  $m$  with  $n$  topological point respectively. Calculate the ratio  $S_i$  of the cumulative distance between the points and the starting point and total length, and put the ratio merged into an ordered sequence  $S = \{s_1, s_2, \dots, s_{m+n}\}$ . In this case, cut out two point in the strokes  $Sa'$  by any two consecutive draw ratios  $S_i$  and  $S_{i+1}$  and suppose  $V_i^{a'}$  is the vector of two points, similarly,  $V_i^{b'}$  is the vector of two points in  $Sb'$ .

The linear similarity  $Sim_{line}$  as in Equation (2) shows: if the  $\alpha(V_i^{a'}, V_i^{b'})$  is negative, the return value is zero and makes the value of  $Sim_{line}$  belonging to  $[0, 1]$ .

$$\begin{cases} Sim_{line} = \sum (s_{i+1} - s_i) \cdot \alpha(V_i^{a'}, V_i^{b'}) \\ \alpha(V_i^{a'}, V_i^{b'}) = \frac{\min \left\{ |V_i^{a'}|, |V_i^{b'}| \right\}}{\max \left\{ |V_i^{a'}|, |V_i^{b'}| \right\}} \cos(V_i^{a'}, V_i^{b'}) \end{cases} \quad (2)$$

Due to the generated direction of the same name strokes being different, it will lead to linear similarity value of zero. Therefore, when calculating the linear similarity, calculate the direction of the maximum similarity as the final direction of strokes by calculating the linear similarity of positive and positive, positive and negative, negative and positive, negative and negative.

Given the road hierarchy of growth structure,  $Sim_{stroke}$  strokes similarity is measured in addition to the linear similarity of strokes, also need to integrate the relative difference of strokes base on the upper strokes to similarity calculation. In the hierarchy it is the set that determines the next layer strokes by searching adjacent strokes of public strokes. So, you can locate a new level strokes by two measures: (1) The relative position of a public stroke to the point of intersection, that is, the cumulative distance length of the point of intersection to the starting point of public stroke; (2) The angle

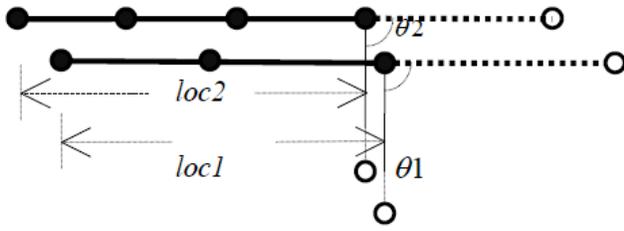


Fig. (4). Relative indicator schematic diagram.

between strokes and public strokes. As shown in formula (3), when two strokes to be calculated are not subject to the initial set to be matched ( $Layer \neq 1$ ), the relative position similarity  $Sim_{loc}$  of the two strokes with the point of intersection in the upper strokes. Direction of the relative similarity  $Sim_{dir}$  base on the upper strokes and linear similarity  $Sim_{line}$  will work together on similarity measure.

$$Sim_{Stroke} = \begin{cases} Sim_{line} & Layer = 1 \\ \varepsilon_1 Sim_{loc} + \varepsilon_2 Sim_{dir} + \varepsilon_3 Sim_{line} & Layer \neq 1 \end{cases} \quad (3)$$

Among them,  $Sim_{loc}$  is the cumulative distance difference of the intersection between strokes to be matched and the starting on the upper strokes, which is the difference between  $loc_1$  and  $loc_2$  in Fig. (4). When two strokes intersect with multiple strokes of the same name in the upper, you need to calculate the average position similarity. Similarly, differences in relative direction  $Sim_{dir}$ , that is, relative deviation angle of two strokes based on the upper strokes, measured by the angle difference between the strokes to be matched and the upper strokes, namely the difference between  $\theta_1$  and  $\theta_2$  in Fig. (4).

Topological similarity of node  $Sim_{node}$  is the basic indicators for measuring similarity between points. Topological similarity of point may be described by the similarity of each adjacent sections, so suppose point  $Pb_j$  and point  $Pa_i$  are the candidate matching points, then the topological similarity  $Sim_{node}^{Pa_i, Pb_j}$  is as shown in Equation (4).

$$Sim_{node}^{Pa_i, Pb_j} = 1/nAE \sum_{AE=1}^{nAE} Max \{ \alpha(AE_{Pa_i}, AE_{Pb_j}) \} \quad (4)$$

$nAE$  is the number of adjacent sections for the point, the formula mode for similarity of sections  $a(AE_{Pa_i}, AE_{Pb_j})$  is as shown in equation (3). In addition, in order to exclude matching directional issues, the bidirectional node topology similarity has to be considered, the average similarity measure formula of the node should be:

$$Sim_{node}^{Pa_i, Pb_j} = 1/nAE \sum_{iAE=1}^{nAE} Max \{ \alpha(AE_{Pa_i}, AE_{Pb_j}) \} \quad (5)$$

$Sim_{node}^{Pa_i, Pb_j}$  is the average node topology similarity of positive  $Pa_i \rightarrow Pb_j$ . Similarly,  $Sim_{node}^{Pb_j, Pa_i}$  is the average node topology similarity of reverse  $Pb_j \rightarrow Pa_i$ .

Define the initial probability  $P^{(0)}$  as the matching degree of node  $Pa_i$  and each candidate point  $Pb_j$ , as shown in equation (6).

$$P_{Pa_i, Pb_j}^{(0)} = \frac{Sim_{Pa_i, Pb_j}}{\sum_{Pb_j \in B_{CS}} Sim_{Pa_i, Pb_j}} \quad (6)$$

## 2) Iterative updating probability matrix.

Iterative updating of the initial probability matrix is to find the best match results from a global point of view, which is the matching process of the point. Probability matrix iterative mechanism is provided by using the formula (7) iteration, defining that exit iteration when the change of two iterative matrixes is less than the threshold.

$$P_{Pa_i, Pb_j}^{(r+1)} = \frac{P_{Pa_i, Pb_j}^{(r)} + S_{Pa_i, Pb_j}^{(r)}}{1 + \sum_{Pb_j \in B_{CS}} S_{Pa_i, Pb_j}^{(r)}} \quad (7)$$

Each iteration needs to be recalculated and the Support Value  $S_{Pa_i, Pb_j}$  of pair to be matched  $Pa_i$  and  $Pb_j$ . Support Value tries to find an adjacent point  $Pb_j^x \in B^x$  of  $Pb_j$  for adjacent point  $Pa_i^x$  in the adjacent point set  $A^x$ , provides a maximum support degree for  $Pa_i$  and  $Pb_j$  using  $Pa_i^x$  and  $Pb_j^x$  matching. Computation formula is as follows:

$$S_{Pa_i, Pb_j}^{(r)} = \frac{1}{nAP} \sum_{Pa_i^x \in A^x} Max_{Pb_j^x \in B^x} \left[ C(Pa_i, Pb_j; Pa_i^x, Pb_j^x) P_{Pa_i^x, Pb_j^x}^{(r)} \right] \quad (8)$$

In equation (8),  $Pa_i$  have  $nAP$  adjacent points,  $Pa_i^x$  and  $Pb_j^x$  is an adjacent point of  $Pa_i$  and  $Pb_j$  respectively. The coefficients of compatible  $Pa_i$ ,  $C(Pa_i, Pb_j; Pa_i^x, Pb_j^x)$  denote compatible degree between adjacent pair point  $Pa_i^x$  and  $Pb_j^x$  and  $Pa_i$  and  $Pb_j$ . If  $Pa_i^x$  and  $Pb_j^x$  are the potential candidate matching point, contribute the support degree of  $Pa_i$  and  $Pb_j$  matching based on matching probability after the  $r$ -th iteration.

The algorithm selects two indicators to measure the compatible degree and to calculate the coefficients of compatibility, direction and length of the road were between adjacent dots.

$$C(Pa_i, Pb_j; Pa_i^x, Pb_j^x) = \sum_{t=(dir, len)} \frac{1}{2(1 + \alpha_t (t_{Pa_i, Pb_j} - t_{Pa_i^x, Pb_j^x})^2)} \quad (9)$$

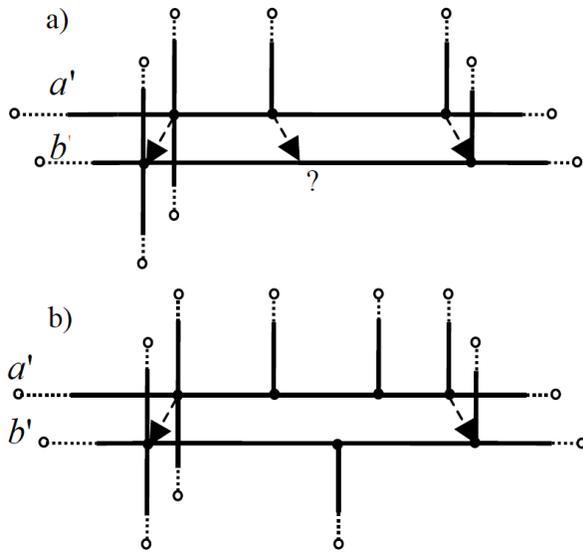


Fig. (5). 1:N,M:N matching condition of nodes.

In order for the compatibility degree to satisfy the requirement of symmetry, make

$$C(Pa_i, Pb_j; Pa_i^*, Pb_j^*) = C(Pa_i^*, Pb_j^*; Pa_i, Pb_j) \tag{10}$$

$$= \sqrt{C(Pa_i, Pb_j; Pa_i^*, Pb_j^*) \cdot C(Pa_i^*, Pb_j^*; Pa_i, Pb_j)}$$

$w_i$  is the weight of the corresponding similarity index. In this paper, it is set to the reciprocal of error factor and the value can be set directly to the global averages, or using Hausdorff distance for estimating.

3) The selection of strokes and the best match relationship for point.

Calculate the sum  $Pn_{Sa', Sb'}$  of maximum matching probability between  $Sa'$  and every point on each candidate strokes  $Sb' \in B_{CS}$ . The corresponding  $Sb'$  with the maximum value of  $Pn_{Sa', Sb'}$  in matching strokes of  $Sa'$ .

After determining the strokes matching relationship, matching point pair of inner strokes should be selected. Steps are as follows:

1) Select the largest similarity measure value from similarity probability matrix between each point and the candidate point, that is matching pair with most matching possibility;

2) Determine the strokes including the two nodes whether matching or not. If not, then keep looking the metric for greater similarity, otherwise the next step;

3) The records of the corresponding point are removed from probability matrix.

Repeat these three steps until no value can be selected. Due to the different number of internal nodes contained in strokes  $Sa'$  and  $Sb'$ , matching the possibility is also different. If the selected point matching relationship is  $1:1$  then the relationships of  $1:N$  and  $N:M$  are deduced by the relationship of  $1:1$ . As shown in Fig. (5), on a pair matching strokes,

if no match occurs between the two points of  $1:1$ , the matching can be inferred as the matching situation of  $1:N$  or  $N:M$ .

The matching method in this paper view nodes as the matching units which are constraint by strokes and the hierarchy. Time complexity of the algorithm is composed by the hierarchical strokes generation, extraction and inner matching. Assume there are  $nm_1$  and  $nm_2$  nodes,  $rm_1$  and  $rm_2$  roads and  $sn_1$  and  $sn_2$  strokes, respectively, in road network data  $A$  and  $B$ . Time complexity of strokes generation and extraction is the linear time complexity for data set reading, writing and seeking operations, the time complexity is  $O(rm_1 + rm_2)$ . Time complexity in inner layer matching has four parts: node topology similarity calculation of initial probability, strokes shape similarity calculation, the support degree calculation of neighboring nodes for iteration and the best results selection. Under ideal state, each node can find only one candidate points in the buffer and time complexity is  $O(nm + sn \times n_{sn} \log n_{sn})$ , therefore,  $nm$  and  $sn$  are the number of nodes and strokes, respectively, and  $n_{sn}$  are the average number of nodes in each strokes. In the worst case, all nodes in the graph to be matched are the candidate to the current node, the time complexity is

$$O(nm_1 \times nm_2 + sn_1 \times sn_2 \times n_{sn} \log n_{sn})$$

Typically, the candidate nodes are much less than total number of nodes, so the time complexity of the algorithm is closer to the time complexity of the ideal state.

#### 4. ALGORITHM ACHIEVED AND RESULT ANALYSIS

We use MAPGIS K9 Software platform with Visual Studio 2010 to implement the algorithm. It can be classified to three stages, data preprocessing, hierarchical iteration and the matching verification.

1) We use geometric network construction methods to build the topology relationship. The process can produce a relation of the arc elements, element nodes, topology and the elements.

2) The generation of layered strokes structure. We describe the roads that have the same average direction as one stroke. In the process of generating stroke, digital direction inconsistency should be tested and reversed to ensure the coordinates are consistent. At the same time, establish the mapping relationship between basic road sections and virtual stroke and also the mapping relationship between virtual path and the internal nodes. Use the road extraction method in literature [19] to extract the first stroke layer, and then generate the adjacent stroke as the next stroke layer.

3) According to the prior knowledge of the two maps such as scale and precision of information, the corresponding point distance deviation is determined as buffer parameter of selecting candidate nodes.

Table 1. Matching results of strokes.

	TP	FP	AM	FN	P	R
Level1	87	1	1	0	97.8%	100%
Level2	121	4	2	8	95.3%	93.8%
Level3	81	2	1	6	96.4%	93.1%
Sum	289	7	4	14	96.3%	95.4%

Table 2. Matching results of nodes.

	TP	FP	AM	FN	P	R
Level1	252	8	2	10	96.2%	96.2%
Level2	541	27	7	20	94.1%	96.4%
Level3	377	26	5	26	92.4%	93.5%
Sum	931	43	14	45	94.3%	95.4%

4) After processing, the original stroke map and target stroke map of first layer are ready. First of all, compute the probability matrix calculation between the initial points. Then, calculate the compatible coefficient and the support degree to get the new probability matrix. When two probability matrixes change less than a specified threshold, the iteration exits. The best match relationships are determined by rules.

5) Process semi-automatic inspection. Check the correctness of the matching relationship. Verification is divided into three parts: 1) The verification of the polysemy node matching results. Polysemy nodes belonging to the same or different multiple road cross layer of nodes, will produce multiple matching relations. If multiple matching relations has agreement, this kind of match will be believed, otherwise the relationship is certain and needs further manual judgment. 2) The detection of weak matching relationship. Detect the similar low probability value of node pairs, to modify the weak matching relationship. 3) The detection of omitted matching. Detecting the nodes which have candidates but ultimately selected for the 1:0 result, manual judgment should be taken to judge whether they are omitted. The three conditions above will be listed, and correct tools are provided.

To verify the validity of the algorithm, a set of Wuhan city road data was selected for validation. The data time of the two matching road data is of the year of 2006 and 2011. According to the metadata, these two data were from different units, as shown in Fig. (7). The base map has 724 roads and 1048 nodes. The map to be illustrated has a total of 1324 roads and 1885 nodes. Test PC has a 2.8 GHz processor and 4 GB of RAM. The value of  $\sigma$  is according to the averages of selected nodes in experimental range. According to the three  $\sigma$  principles radius of buffer parameter is set to 200 meters. The terminating threshold of probability matrix was set to 0.005 and the single iteration of the matching process takes an average of 24.6 s.

Four indicators were selected to evaluate the matching accuracy: Really optimistic matching TP (The essence of

correct matching), False optimism matching FP (The essence of mismatch), False negative match FN (The elements that should be matched but does not), Not sure match AM (Artificial uncertain matching elements). Compared with the manual matching results, the result of this matching algorithm in this paper is shown in Table 1 and 2.

The formula of the matching accuracy P and the recall rate R is:

$$P = \frac{TP}{TP + FP + AM} \times 100\% \tag{11}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{12}$$

Level 1-3 of Table 1 shows that the matching result of the three-level strokes, the row named sum represents the overall matching result. The layer index and the overall index are all calculated according to the formula 11 and 12.

In this case, the accuracy rate of point matching is 94.3%, recall rate is 95.4%. There are 4.6% of matched pairs that could not be effectively identified. The matching result is more than satisfactory. As Table 1 and 2 show, the accuracy of matching results of strokes is slightly higher than the node. This is because the strokes have higher identification ability than a single node. And whether the nodes or strokes, the matching accuracy of Level 1 is higher than Level 2-3. The matching relationship of senior network features are easier to determine than low-level network features. Compared with the literature [18] which used the probabilistic relaxation method and applying the point as the basic matching unit, the algorithm in this article integrated the constraints of strokes and road level, so that the algorithm improved the matching accuracy to some extent.

Fig. (6) shows the matching results of the first layer of the strokes. In Fig. (6a), there is a situation of 1:0 in strokes in the presence of the two road network data, such as the red

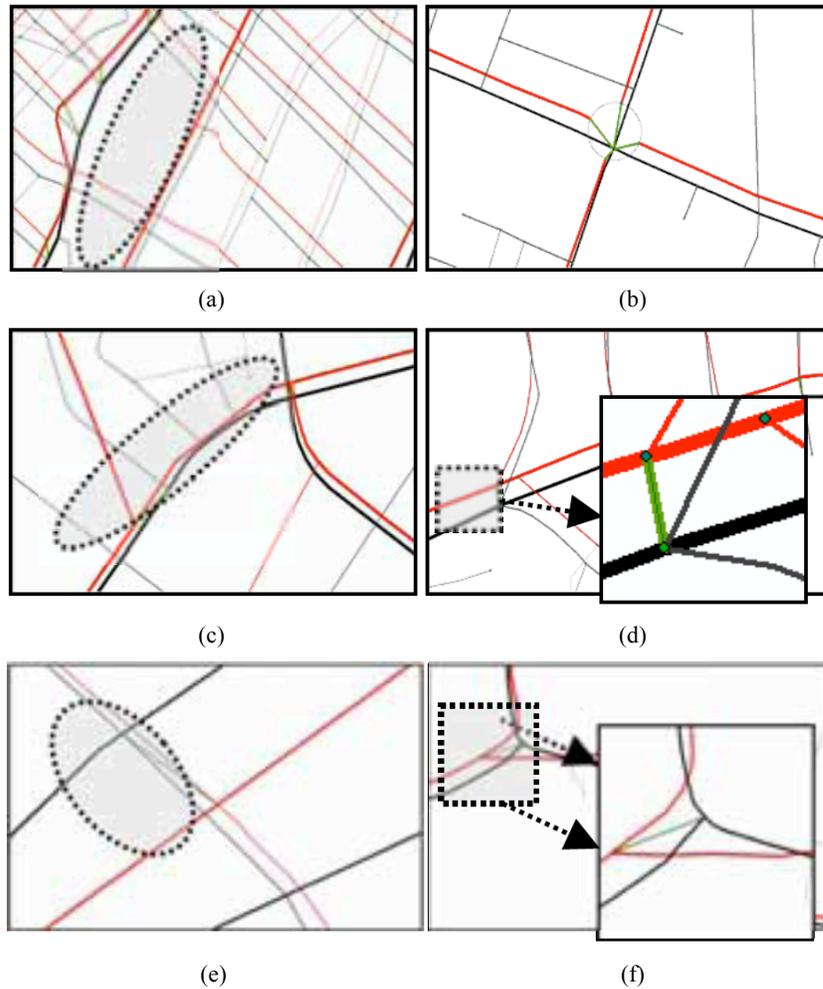


Fig. (6). Matching result of first layer.

strokes in dashed box, if stroke was not matched in first layer, it will be downgraded to continue participating in a matching in the next layer eventually. And there is an uneven deviation in the left side of the strokes, the correct matching could be found from nodes of strokes. In Fig. (6b), there are two different expressions about the intersection of the two road network data. The red road network is a ring junction and the black road network is the crossroads. The method in this article is to disconnect in the ring junction when generating the strokes, retain the structure of ring junction when matching, so as to avoid tedious streamline operations and find the correct matching relationship. In Fig. (6c), there is a situation of  $1:0$  in strokes, the first layer black strokes connect with three level 2 strokes in dashed box, generating three nodes, and no node corresponds to the red strokes, so by correcting the two ends corresponding node of  $1:1$ , the correct  $1:N$  matching relation is obtained when matching. As shown in Fig. (6d), when the node is determined to match the first layer  $1:1$  matching node and in fact the node is a  $1:N$  matching node, this problem will be resolved in the lower matching.

Fig. (7) shows the matching result of the second layer. The most network strokes are in the second layer and the

nodes of strokes will be constrained by the structure of the multilayer strokes, the matching relationship will appear in multilayer simultaneously. The three level matching relationship, as shown in Fig. (7a) demonstrates that when the ambiguity or polysemy exist, we should compare the results of different layers, if consistent, accept it, if not, we require judging by ourselves. In Fig. (7), the part of the strokes connected with the intersection was selected as the first layer strokes, the rest was selected as the second layer strokes, so there are four endpoints which were matched in the first layer which is the yellow line, and two endpoints were matched in the second layer which is the blue line. Fig. (7c) shows situation of  $N:M$  in the second layer, the same, we should determine the matched pair of  $1:1$  in strokes, getting matching result of  $N:M$  of roads. Fig. (7d) shows that the node do not find the correct relations, same as Fig. (6d), however two endpoints belong to two strokes in this layer and getting matching result of  $1:N$ . Fig. (8) shows the matching results of the third layer strokes. The third layer contains fewer number of strokes, formed by the third layer strokes (Fig. 8a, b) and the strokes (Fig. 8b) which are not matched, are downgrade.

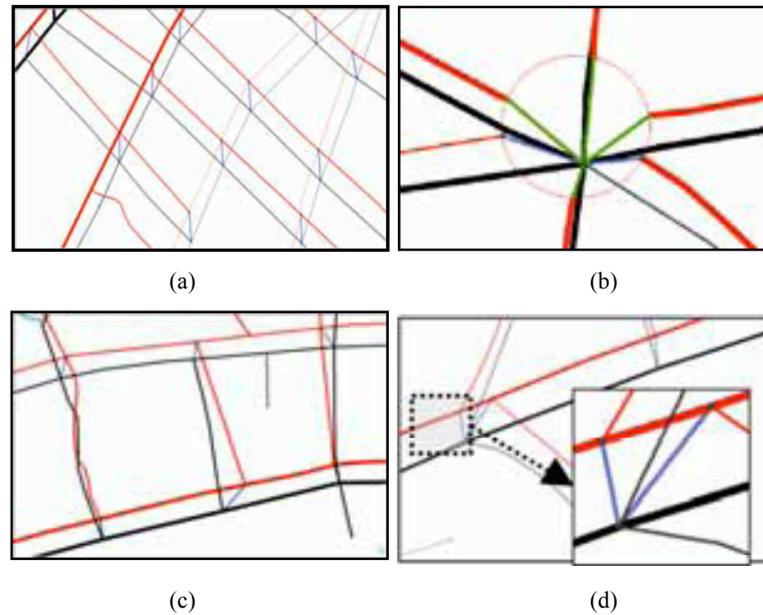


Fig. (7). Matching result of the second layer.

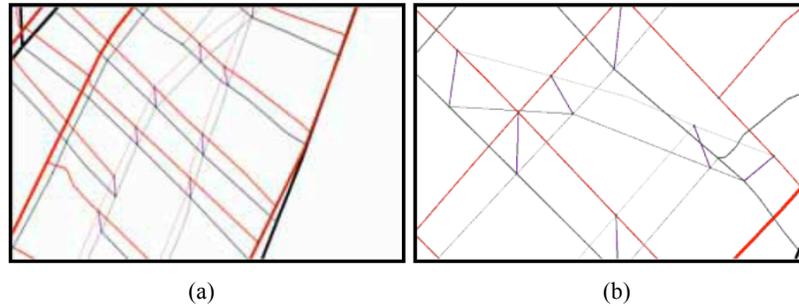


Fig. (8). Matching result of the third layer.

**CONCLUSION**

This paper presents a road network matching method based on hierarchical stable strokes, the characteristics of the method is that the road network level concept combined with a cross road structure, adopts the top-down hierarchical matching strategy combing with probabilistic relaxation method. This strokes information provides a common reference for matching, to facilitate the conversion and establishment of computation space, at the same time, provides more context information for the matching relationship in determining between the partial elements. Constrained by multi-layer strokes improve the ability to identify matching relations. This method can not only solve the matching problem in the conventional case but also is better able to handle  $1 : N$  and  $N : M$  matching situations, overcoming individual similarity measure partial failures in local matching method. Shortcomings of this method are that: 1) the length of strokes have influence on the relative position indicator, therefore, it needs to control the length of strokes, using the principle of complex traffic mouth truncation to ensure the relative position indicator sensitivity; 2) When determining the matching complexity, the matching probability between points  $1 : 0$  in the probability calculation formula can be added, thereby reducing mismatch. Following research will focus on improving efficiency of the algorithm in order to meet the matching needs of large-scale road data.

**CONFLICT OF INTEREST**

Financial contributions to the work being reported should be clearly acknowledged, as should any potential conflict of interest.

**ACKNOWLEDGEMENTS**

This research has been supported by National Natural Science Foundation of China (No, 41201385, 41301426, 41301427, 41371422), Open fund of State Key Laboratory of Geographic Information Engineering (SKLGIE 2013-M-4-5) and the Fundamental Research Funds for the Central Universities (CUGL120274).

**REFERENCES**

- [1] J. Juan, F. Javier, and A. Manuel, "Digital map conflation: a review of the process and a proposal for classification", *International Journal of Geographical Information Science*, vol. 25, pp. 1439-1466, Sept., 2011.
- [2] B.B. Zhao, M. Deng, and X. Xu, "Development of general rules for matching multi-scale area objects", *Geomatics and Information Science of Wuhan University*, vol. 36, pp. 991-994, Aug., 2011.
- [3] B. B. Zhao, "A study on multi-scale vector map objects matching method and its application," Ph. D. thesis, Central South University, Changsha, China. 2011.

- [4] M. Deng, B. B. Zhao, and Z. Xu, "Representation methods of distance between spatial objects in GIS and their analysis", *Computer Engineering and Applications*, vol. 47, pp.35-39+45, Jan, 2011.
- [5] Y. Gabay, and Y. Doytsher, "Adjustment of line maps", in Proceedings of GIS/LIS'94, *Phoenix Arizona*, pp. 333-341, 1994.
- [6] P.C. Liu, J. Luo, and T.H. Ai, "Evaluation model for similarity based on curve generalization", *Geomatics and Information Science of Wuhan University*, vol. 37, pp.114-117, Jan., 2011.
- [7] M.Deng, M.L.Niu, and Z.L. Li, "A generalized hausdorff distance for spatial objects in GIS", *Geomatics and Information Science of Wuhan University*, vol. 32, pp.641-645, Jan., 2007.
- [8] C.B. Jones, D.B. Kidner, and L.Q. Luo, "Database design for a multi-scale spatial information system", *International Journal of Geographic Information Systems*, vol. 10, pp. 901-920, Aug., 1996.
- [9] M.Sebastain, and D.Thomas, "Matching networks with different levels of detail", *Geoinformatica*, vol. 12, pp. 435-453, Dec., 2008.
- [10] Y.M. Chen, J.Y. Go, and W. Z. Shi, "A distance-based matching algorithm for multi-scale road networks", *Acta Geodaetica et Cartographica Sinica*, vol. 36, pp. 84-90, Jan., 2007.
- [11] Y.H.Wang, "A spatial entity matching approach for updates extraction and integration", *Science of Surveying and Mapping*, vol. 36, pp. 128-130, Jan., 2011.
- [12] E. Safra, Y. Kanza, and Y. Sagiv, "Ad hoc matching of vectorial road networks", *International Journal of Geographical Information Science*, vol. 27, pp. 114-153, Jan., 2013.
- [13] D.B. Zhao, and Y.H. Sheng. "Research on automatic matching of vector road networks based on global optimization", *Acta Geodaetica et Cartographica Sinica*, vol. 39, pp. 416-421, Apr., 2010.
- [14] B.S. Yang, Y.F. Zhang, and X. C. Luan., "A probabilistic relaxation approach for matching road networks", *International Journal of Geographical Information Science*, vol. 27, pp. 319-338, Feb., 2012.
- [15] Y.F. Zhang, B.S. Yang, and X. C. Luan, "Automated matching urban road networks using probabilistic relaxation", *Acta Geodaetica et Cartographica Sinica*, vol. 41, pp. 933-939, Jun. 2012.
- [16] M. Zhang, and L. Meng, "An iterative road-matching approach for the integration of postal data computers", *Environment and Urban Systems*, vol. 31, pp. 597-615, May, 2007.
- [17] Y.G. Hu, J. Chevn, and R.L. Zhao, "Matching of road under different scales for updating map data", *Geomatics and Information Science of Wuhan University*, vol. 35, pp. 451-456, Apr., 2010.
- [18] W.B. Song, M. James, and L. Timothy, "Relaxation-based point feature matching for vector map conflation", *Transactions in GIS*, vol. 15, pp.43-60, Jan., 2011.
- [19] B.S. Yang, X.C. Luan, and Q.Q. Li, "Generating hierarchical strokes from urban street networks based on spatial pattern recognition", *International Journal of Geographical Information Science*, vol. 25, pp. 2025-2050 Dec., 2011.

---

Received: September 22, 2014

Revised: November 05, 2014

Accepted: November 06, 2014

© Yang et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.