A HOPFIED NEURAL NETWORK ALGORITHM FOR AUTOMATED NAME PLACEMENT FOR POINT FEATURE

Fan Hong Zhang zuxun Du Daosheng
The State key Lab. of ,WTUSM,39 Luoyu Road,Wuhan,China,430070
Tel: 86-027-87881292, Fax:86-027-87643969
fh@hp01.wtusm.edu.cn

KEY WORDS: Hopfield neural network, Energy function, Map name placement.

ABSTRACT

This paper presents a method of adding label to the map especially for the point feature. This method overcomes the shortcoming of traditional methods eg. Conflict-backtracking method. Its kernel algorithm use the hopfileld neural network to find the best label position for point feature. The experimental results proves that this algorithm has good permanence and high speed.

1 PREFACE

Map name is an important component of the map, whether labels in place or not plays an important role on the readability and usefulness of a map. Thus far, the placement of cartographic names has been a manual process which takes plenty of time and energy. More recently, many different algorithms have been developed to aid this process. Among them are the priorities suggested by Yoeli[Yoeli ,1972] for the placement of names for point feature, the integer programming is designed for the placement of names for point and linear feature[Zoraster, 1986] and an expert system approach for dense-map name placement[Doerschler, 1989]. Yet the automation of name placement has proved difficult to be implemented successfully. Because an automated name-placement system must select names from scale-independent databases and place them in a manner acceptably similar to that of a trained cartographer. Latest research shows that the kernel problem of map name placement is a NP-hard problem, one way to solve this problem is to try some way to decrease the complexity or improve the efficiency of problem in order to find the near-optimal position in the user-tolerable time .

This paper regards this problem of map name placement as a typical combinatorial-optimal problem and propose to makes use of the neural network methods to find the best name position of local or global for each map resident feature. Compared with those of the traditional conflict-tracking methods, this method through experiment turns out to be better and more efficient.

2 EXPERIMENTAL DATA

Our experimental data comes from the 1:25 National Base Map Database and comprises 3 topographical map sheets that are stored in the vector format of ARC/INFO. The map-no of the three maps are respectively H-48-[10],H-48-[13] and H-48-[14], each of which is stored into 16 layers that include hydrogen, roads, vegetation and administrative boundary. These data contain both the spatial features in the form of digital map and the attributes information such as map name and the feature code in terms of the national standard encoding of geographic feature by which we can easily discern label attributes such as size, style and spacing etc.

The 3 map sheets contain respectively 3251,1651 and 2734 resident points. The area of these maps is near Chendu city , the capital of the Sichuan Province in the west china and is well-known for its dense-population. Because map name placement in a dense map is more difficult than in a sparse map, these 3 maps among our experiment data can well serve as good representatives.

3 NAME PLACEMENT PRINCIPLE FOR POINT FEATURE

Map features can be categorized into points, lines and polygons in terms of their spatial distribution, and name placement of the three kinds of geographic features will conform to different cartographic conventions. However, after studying various kinds of maps and labeling procedures , we find the following principles is in common adaptable.

(1)"belong to" principle

Label should refer unambiguously to its intended referent feature. Namely the relation of the label and the referent feature should be easily recognized but not be confused with the other nearby labels and other geographic feature.

(2)"avoiding off" principle

The label name should make away for important geographic features, but not overlap important features, especially the features or labels of the same color.

(3)"accustomed to" principle

The character position、character order or so on should accord with the reading custom of the readers. To Chinese readers reading permits some flexibility, not rigidly adhering to a left-to-right scanning of words. Both vertical and horizontal alignments are commonly used.

The point name placement algorithm is used to identify point feature depicted by pictorial symbols. The point symbol is centered on a single coordinate. According to both the principle of the above mentioned and the particularity of the point feature, the principle suitable for point feature was given as follows.

(1)According to the map-making regulations of 1:250000 topographic of the People's Republic of China, the procedure for automated name placement is developed. One of the eight ranked positions surrounding a feature symbol with referencing fixed positional weight is used to place a point feature label. In order of preference, these positions are : to the right of the symbol,above,to the left, below, above and to the right, above and to the left,below and to the left , below and to the right, showed in Figure 1.Position with higher weights are preferred over positions with lower weights.

(2)Point label should not overlap the placed label and any point feature.

(3)Point label should not overlap the important linear feature of the same color such as railways and major roads etc, While overlap is unavoidable, efforts should be made to decrease it.

(4)labels and labels must not overlap each other.

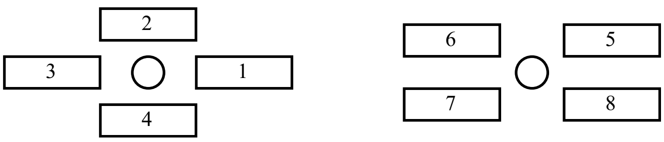

Figure 1. Ranked label positions for point feature

(5)Label name and the referent features had better locate in the same side of the nearby linear feature and label name can not overlap boundary.

4 THE SOLUTION TO THE AUTOMATED NAME PLACEMENT FOR RESIDENT FEATURE

Point feature involves resident、flag point and elevation point etc., among which resident is representative for its largest number and highest density , According to the principle of name placement for point feature mentioned above , the following strategies are taken to automate name placement for resident of topographical map.
The total solutions contain the following 3 stages:

step 1: coarsely choose the candidate position and determine the optimal level for every position.

First of all, 8 available position for every resident are chosen, showed in Figure 1, through evaluation of "readability" and "belong to" relations of the 8 positions , a weight factor is given for every possible position and detailed method is depicted as follow.

The weight of the candidate position is the function of the primary weight and the secondary weight, that is :
weight = the primary weight * the secondary weight
Specify a basic weight from 1.0 to 0.93 corresponding to the priority from high to low.
Use a rectangle area to represent the area that a label name occupies and judge if the area will overlap the important feature such as railway, major road and minor road. Different secondary weight factor is assigned accordingly. For example, if the label rectangle of one feature overlaps with a railway or a major road , according to the extent of overlap, its secondary weight should be assigned to 0.1, 0.2, 0.3 or 0.4 respectively; In order to control overlap with the minor road , the secondary weight of 0.51,0.52,0.53 should be assigned for different extent of overlap.
Judge whether the label and the referent feature appears on the same side of the boundary and whether the label overlap with the boundary, To handle this case, an appropriate secondary factor should be designated.

Step 2: construct a hopfield network, set the initial value for it and make it run. If necessary, make the network run many times, observe its convergence, select the best results and record it.

Step 3: after taking the previous steps, local optimal processing is made so as to resolve the remaining conflicts between the residents labels. Few remained conflicts that have not yet been resolved in the end, will be adjusted and revised manually.

The core of the whole algorithm is step 2, that is to use hopfield network to find the best local or global label position for every resident on topographical map.

This hopfield approach runs the network in an iterative way that will converge very fast avoiding of many times backtracking and the nested backtracking of the traditional method, which will, as a result, improves the efficiency of the search algorithm.

5 THE ALGORITHM OF THE HOPFIELD NEURAL NETWORK TO FIND THE BEST LABEL POSITION FOR RESIDENT

Hopfield neural network is a one-level of feedback network , Let N1, N2, …Nn stand for the n neural unit, Wij stands for the connection weight from Ni to Nj. If we use W to represent the connection strength between the n nodes, Hopfield is symmetric, then:

$$W_{ij} = W_{ji} \qquad i, j \in \{1, 2, 3 \ldots n\}$$

For the continuing feedback network, when the network is working, the relation between the input and the output can be represented in terms of the following status equation , among which, g(。) is a continuing monotony ascending function with up limit, among which sigmond or hyperbolic tangent function most commonly used, Ui represents the input of the neural unit i , Vi represents the output of the neural unit I, Ii represents the bias of the neural network namely the stimulus coming from the external world.

$$C_i \frac{du_i}{dt} = -\frac{u_i}{t} + \sum_j W_{ji} V_j + I_i$$

$$V_i = g(U_i)$$

If the evolution of the equation takes the asynchronous way, at any time , only one status of one neural unit will be changed. Suppose Ui(t) stands for the sum of all the inputs of the neural unit i at the time of t, Vi(t+1) stands for the output status of this neural unit at the time of t+1,then:

$$U_i(t) = \sum_j W_{ji} V_j(t) + I_i$$

$$V_i(t+1) = g(U_i(t)) = g(\sum_j W_{ji} V_j(t) + I_i(t))$$

For the hopfield network, the Lyapunov energy function will take the following form:

$$E=-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}T_{ij}V_iV_j-\sum_{i=1}^{n}V_iI_i$$

$$U_i(t)=\sum_{j}W_{ji}V_j(t)+I_i$$

$$V_i(t+1)=g(U_i(t))=g(\sum_{j}W_{ji}V_j(t)+I_i(t))$$

$$E=-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}T_{ij}V_iV_j-\sum_{i=1}^{n}V_iI_i$$

It has been proved that hopfield network is a non-linear motion system. There exists one or more minimum point or balancing point. At some time , after the status of every neural unit is given and the initial network status is set up , the status of the network will change in the direction of energy gradually decreasing in the light of the working equation, and in the end, approach or reach the balanced status of the network, which is the minimum point of the energy. This is convergence of the energy of the hopfield neural network. In this way, while the energy function is converged to a minimum point, the best solution for the problem will be produced.

The problem of finding the best label position for resident feature can be regarded as a combinatorial optimal problem , A hypothesized map contains m residents, and every resident has n candidate points(such as, n = 8, 16…). There are m ×n candidate points , and these residents will be listed as a matrix in which n candidate label positions of one resident will be lined as one row, altogether there are m row * n column, as showed in table 1. Considering this table as a matrix of m rows * n columns, m stands for the number of the residents, n stands for the candidate label position for every resident. If we correspond one candidate label position to one neural unit, thus we can construct a hopfield neural network that comprises m×n neural unit.

In order to define the energy function , we describe the problem as the the sum of constraint condition and the optimal goals as follows:

	candidate label position							
Resident point	1#	2#	3#	4#	5#	6#	7#	8#
1#	1	0	0	0	0	0	0	0
2#	0	1	0	0	0	0	0	0
3#	0	0	1	0	0	0	0	0
...	...							
...	...							
2514#	1	0	0	0	0	0	0	0
2515#	0	1	0	0	0	0	0	0

Table 1 candidate label position for resident

Constraint condition: every resident can choose only one label position.

The most optimal goal: the number of overlapping between two label rectangle area is the smallest.
According to the above constraint condition and the most optimal goal, we can give out the energy function of the network as follows.

$$E=\frac{B}{2}\sum_{i}\left(\left|\sum_{j}V_{ij}\right|-1\right)^2+\frac{A}{2}\sum_{i}\sum_{j}\sum_{k}\sum_{l}D(i,j,k,l)V_{ij}V_{kl}$$

Of the above equation, the first item represents the constraint condition, and only one label position is allowed for one resident feature. When this condition is met, the first item will be equal to 0. The second item is optimal goal, D(i,j,k,l) is specified as follow:

$$D(i,j,k,l)=1 \quad when V_{ij,} V_{kl} \quad overlap \quad each \quad over$$
$$0 \quad when V_{ij,} V_{kl} \quad do \quad not \quad overlap \quad each \quad other$$

Thus the second item is the multiplied number of the overlap between every two label rectangles, if the choice of label position is the most optimal, E_2 can reach the minimum value, and if the choice of label position is more optimal, E_2 can reach the smaller value.

By comparing the energy function with the standard energy function, the connecting weight between unit i and unit j will be determined as:

$$T_{ij,kl} = -AD(i,j,k,l) - B\delta_{ij,kl} \quad ......(1)$$

$$\delta_{ij,kl} = \begin{matrix} 1 & \text{当}ij=kl\text{时} \\ 0 & ij \neq kl\text{时} \end{matrix}$$

$$I_{ij} = B \qquad ......(2)$$

Because $W_{ij,kl} \propto T_{ij,kl}$,let$W_{ij,kl} = T_{ij,kl}$, then the running equation will be derived as follow:

$$\frac{du_{ij}}{dt} = \frac{-u_{ij}}{\tau_i} + \underset{j\neq i}{\quad} \left(AD(i,j,k,l) - B\delta_{ij,kl}\right)V_{ij} + B \qquad ......(3)$$

$$V_{ij} = g(u_{ij}) \qquad ......(4)$$

Here we adopt sigmoid function as the I/O function of the neural unit, and ij is the subscript ,which stands for the neural unit of jth label position of the ith resident.

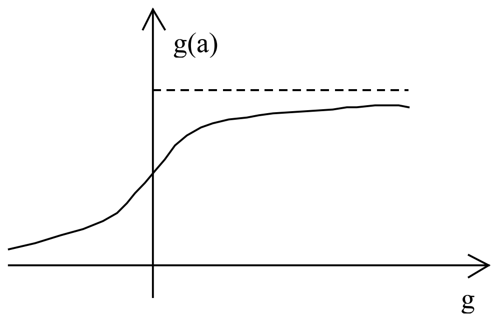

Figure 2. sigmoid function

The detailed calculation procedure and the iterative step are listed below:

(1)The initial value is set up in the light of the above guideline；

(2)Calculate the output of every neural unit $V_{ij}(t_0)$ according to $V_{ij} = g(u_{ij})$ ；

(3)Put $V_{ij}(t_0)$ into(2),calculate $\frac{du_{ij}}{dt}|t = t_0$ ；

(4)According to equation $u_{ij}(t_0 + \Delta t) = u_{ij}(t_0) + \frac{du_{ij}}{dt}|t = t_0 \Delta t$, calculate $u_{ij}(t + \Delta t)$ of the next time point.

(5)Return to step (2).

6 V EXPERIMENTAL RESULTS AND COCLUSION

Experience of running the equation shows that while an appropriate parameter and an appropriate constant of the time τ_i are specified, the network will converge normally to a satisfactory end status.

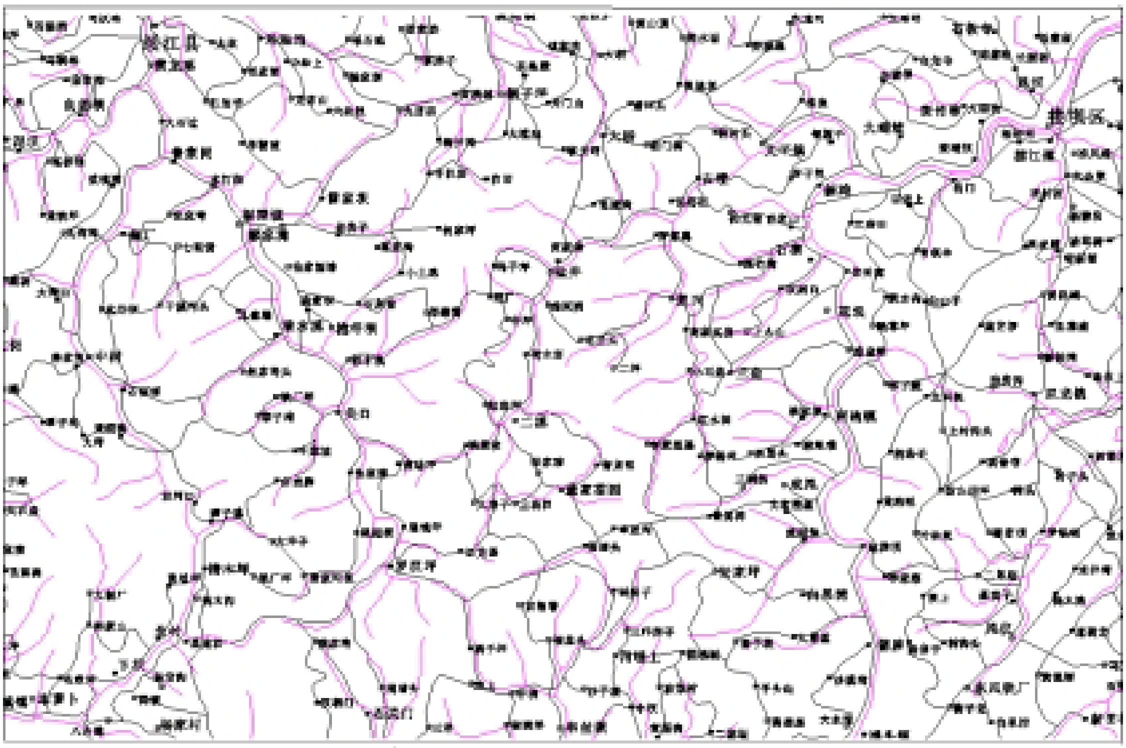

Figure 3. Results of automated name placement for resident feature

Our original spatial data can not be directly used in our algorithm. These map data must take preprocessing before being available. The preprocessing operations consist of the symbolization of map feature , the conversion of vector feature to raster format, the coding of feature and the overlay of the raster maps. Besides, raster spatial information will be written into different binary files in terms of their feature type. Meanwhile, the attribute information of the name will be read out of the source file and be written into a binary file to be easily used

The label position for the residents, after 10 times of iteration of the neural network, we can reduce the conflict number The 3 maps comprise 3251, 1651 and 2734 residents features respectively. Making use of the above algorithm to find to 196, 97 and 123. After the local optimization and adjustment, all conflicts are resolved, the running time will be respectively 19, 11 and 15 minutes. The result is satisfactory. Figure 3 shows the result of adding resident label to h-48-[10] with this algorithm .

The experimental results show that the neural network method can be used to solve the combinatorial optimal problem including finding the best label position. Because the network can converge swiftly, therefore complex problem with "combinatorial explosion" danger can be converted to a simple problem, which will largely improve the efficiency of algorithm.

The problems of name placement differ significantly among the three categories. The approach to the placement of names for points is relatively easier than linear and areal features. In our experiment, we merely consider resident feature so that the problem can be simplified. Now we are doing further research to make our algorithm more powerful so as to solve more problems and suit more cases.

REFERNENCES

Hiersch,S.A.,1982,“An Algorithm for Automatic Name Placement Around Point Data.” The American Cartographer, vol 9., No. 1, pp 5-17.

Ann,J. And Freeman.H.,1983,“A Program for Automatic Name Placement”, Proceedings of Auto-Carto VI, Vol. 2, pp. 444-453.

Lee R.Ebinger and M. Goulette, 1990 ,“Non-interactive Automated Names Placement for the 1990 Decenial Census.”, cartography and Geographic Information systems, Vol. 17, No.1, pp. 69-78

Doerschler,J.,and H.Freeman,1989,“An Expert System for Dense-Map Name Placement.” ,Proceedings of Auto-Carto 9,pp. 215-224.

Steven Zoraster,1986, “Interger Programming Applied to the Map Label Placement Problem.”, Cartographica , Vol.23, No.3, pp16-27.

Cormely,R.G.,1985,“An LP Relaxation Procedure for Annotation Point Feature Using Interactive Graphics.” Proceedings of Auto-Carto 7,pp. 127-132.

Dacid S. Johnson, Umit Basoglu,1989, “The Use of Artificial Intelligence in The Automated Placement of Cartographic Names”, Proceedings of Auto-Carto 9.

Freeman,H., and J. Ahn,1984, “AUTOMAP: An Expert System for Automatic Map Name Placement.” ,Proceedings of the First International Symposium on Spatial data Handling, pp 544-571.

Marks J, Shiebar S.,,1991, “The Computational Complexity of Cartographic Label placement ”, Technical Report , Harvard University.

Yoeli,P.,1972, “The logic of Automated Map Lettering.” ,The cartographic Journal, vol. 9,n.2,pp.99-108.

James E. Mower,1993,“Automated Feature and Name Placement On Parallel Computers”, Cartography and Geographic Information Systems, Vol.20, No.2, pp. 9-82.

Zoraster,S.,1991,“Expert Systems and the Map Label Placement Problem.” Cartographica, vol. 28, no. 1, pp. 1-9.