Graph Partitioning Method to Divide Rural Financial Information Service Area

Han Dong-mei
Harbin University of Commerce
Harbin, China
e-mail: handm@hrbcu.edu.cn

Zhou An-yu
Harbin University of Commerce
Harbin, China
e-mail: zhouay@hrbcu.edu.cn

Yao Feng-ge
Harbin University of Commerce
Harbin, China
e-mail: yaofg@hrbcu.edu.cn

Wen Hong-mei
Harbin University of Commerce
Harbin, China
e-mail: wenhm@hrbcu.edu.cn

Dong Xiao-hong
Harbin University of Commerce
Harbin, China
e-mail: dongxh@hrbcu.edu.cn

Wang Shu-yan
Harbin University of Commerce
Harbin, China
e-mail: wangsy@hrbcu.edu.cn

Han Song-yan
Harbin University of Commerce
Harbin, China
e-mail: hansy@hrbcu.edu.cn

Zhang Bing
Harbin University of Commerce
Harbin, China
e-mail: zhangb@hrbcu.edu.cn

Song Peiyou
The University of New Mexico
Albuquerque, NM USA
e-mail: peiyou@cs.unm.edu

Shu Anhei
Rice University
Houston, TX USA
e-mail: anhei.shu@rice.edu

Abstract—Divide rural financial information service area divided using graph partitioning methods. By means of one-to-one mapping, the factors of the activity place (area) for rural financial information service object, including population, area and financial activity quantity are converted to corresponding road intersection (nodes), according to certain rules, in roads as edges connecting each node, and the basic topology of the service area is obtained. Using multilevel k-way partitioning algorithm, the nodes are divided into different partitions, and the division of the service area is implemented. The data mapping conversion tool is given, the data file format is specified; implementation of graph partitioning algorithm have been elucidated; the evaluation function is constructed. Graph partitioning method for rural financial information service area divided both to ensure a balanced partition of the service area, but also to ensure the connectivity of partition interior; for the optimization of rural financial information network layout, provides a scientific decision-making tools; applications of graph partitioning method is extended.

Keywords- finance; graph partition; information; service area; indicator

I. INTRODUCTION

Rural financial information undertakes the historic mission to improve rural financial services quality, to promote coordinated development of finance and economy both in cities and countryside, and to push to build a new socialist countryside and harmonious society, which plays an important role to get a communication channel between the government and farmers, financial institutions and farmers, to improve financial environment in countryside, to promote agricultural industrialization management, to implement national benefitting farming policies, to maintain social stability and economic development in countryside. But in rural financial information development, the problem is that service points unreasonably layout. namely service area unreasonably division. Affected by various factors, such as geographical environment, traffic conditions, historical conditions and local finance etc, the density of rural financial service points is much lower than it in cities. Service points density in developed areas is large, service points are less in new residential and remote areas. Rural financial information service points with all function are scarcer. It causes inconvenience to financial activities of some farmers, seriously affected the farmers' financial enthusiasm activities. Studies shows that lack of financial organization is one of the main reasons that the rural financial institutions are inefficient. In order to change it is necessary to reasonably divide rural financial information service areas, and to optimize the service points distribution.

Graph partitioning is an important research method, which is widely used in VLSI design, parallel computing, task scheduling and other fields. When the tasks should be divided into multiple parallel parts, and each part of the workload should be equal as far as possible. The role of graph partitioning method is very outstanding. Such as: task allocation and partitioning problem in heterogeneous
computing environments [8, 9], dynamic load balancing problem [10-12], etc. In recent years, with the development of complex network theory, graph partitioning is also applied to the social network research [13]. Because of the need to research Large-scale network simulation [14, 15] and cognitive networks [16], it is deepening to study that a network partitioning problem is abstracted into a network topology graph partitioning problem. However, it’s hard to find rural financial information service areas partitioning by graph partitioning method. It can be imagine that if according to certain rules, financial service objects are abstracted as nodes, then all of the service objects are abstracted into a node set graph. Thus, using graph partitioning method according to certain rules service area can be partitioned.

II. FACTORS AFFECTING THE RURAL FINANCIAL INFORMATION SERVICE AREA DIVISION

The division of the financial services area is the base of financial services center outlets setting. There is a positive correlation between the numbers of financial service center outlets set with the economic development level of the researched area. It is should be based on the level of economic development of different areas, planning the number of rural financial service center outlets settings and the construction speed [18]. On the basis of determined number of rural financial services center outlets set, to determine how its distribution, that is division of the corresponding service area, it is the key to bring into play the financial services center functions, improve the service efficiency and expand the service area. There are many factors which influence division of the rural financial Information services area, it is needed from the numerous factors of macro and micro, economic and geographical, to carry out logical analysis and experience summary, refined out the main influencing factors which have Strong objectivity common view, comparative, quantization, and maneuverability.

Geographic information system provides an effective tool for analysis and identifying the related influencing factors. By geographic information system for the spatial data collection, storage, management, analysis, the object of representing and simulating real world space, analysis and processing spatial data can be achieved. Using this tool, Liu Zong-wei [19], Huang Ling [20], etc. put forward a comprehensive site selection method based on GIS and neural network; Gong Yu-xia, Wu Yu-hua put forward triangular fuzzy number comprehensive evaluation method for site selection [21]; depending on the layout status of population, traffic, commercial and financial outlets in the service area, using GIS technology, Yang Liu, Tian Yongzhong, Wang Jie-chen provided a scientific basis for the ATM machine outlets location decision [22]. In the case of rural banks of Jiangsu province, Zhang Bing, Li Dan analyzed the influence factors of new rural financial institutions outlets layout, Put forward the main influence factors of four classes: Demographic factors, economic activity, financial infrastructure status and the financial institutions themselves competitive factors [23]. Liu Yanfang, Gu Xing, Li Chun-yan, Jin Hai, etc. based on autocorrelation analysis for Local space between a ATM and more features pointed out that population density, road density, finance, business and density of public facilities, etc. affect the spatial layout of the ATM on the global and local [24]. Comprehensive analysis of the literature in many aspects, this paper deems that the operational, quantifiable main factors play a decisive role to the rural finance information service area division is: population, area proportion, the annual per capita financial activity quantity, etc.

III. GRAPH PARTITIONING METHOD TO DIVIDE RURAL FINANCIAL INFORMATION SERVICE AREA

Rural financial information service area division is similar to the task allocation problem, by financial institutions of the corresponding area to provide services for the service object, is equivalent to the service object is divided into a corresponding financial services area. The rural financial information service area division involves to an influence factors of various ways, objects and environments, etc. of information service, is limited by a variety of constraints, so it is a more difficult problem. Using graph partitioning method can deal with the partitioning problem with a variety of attributes, which has different weights each, partitioning condition with multiple constraints. Therefore, after the influence factors of rural financial information services area division is determined, according to certain rules, rural financial information services area division is transformed into graph partitioning problems, on this basis, determining the partitioning objectives and constraint conditions should be satisfied, using an appropriate partitioning method to partition, then the results of the optimization partitioning the service area can be derived.

A. The transformation of the service area partitioning problem towards the graph partitioning problem

Graph partitioning problem is according to certain requirements and rules to partitioning a graph into roughly equal several partitions. The graph partitioning method is introduced into the rural finance information services area division, need to adopt means of one-to-one mapping, the rural financial information service object activity place is transformed into the nodes in the graph according to certain rules, the influence factors of rural financial information service area division is transformed into the node weights of different attributes, thus the original service area data graph is transformed into the basic topology graph. On this basis, according to the basic principle and method of graph partition the nodes are divided into different partitions, thus the service area division is realized.

Generally speaking, the basic unit of rural financial information service object should be the natural person (or corporation), every natural person (or corporation) should be a node in a graph, but such data volume is too big, and the key is impossible at all specific to master each and every one. However it can do that according to macroeconomic statistics to understand or estimate the general situation of a service area. Therefore, in order to take advantage of graph partitioning method, the live place or workplace (such as residential, office buildings) of service object (natural person or corporation) is abstracted into service area. Each service area, according to its population, area, the annual per capita financial activity quantity differ, having the different weights of a plurality
attributes. Through its geographic information, the area of the service area can be calculated. According to the orientation of its exports, plurality attributes weight of each service area are statistics to the right front intersection (node) of close the road, as shown in Figure 1. The node weight is the sum of weights of each service area corresponding attributes moved to the node. Nodes connected by edges (roads).

In Fig. 1 service area a has 3000 people, 34.56 square kilometers, AFAQPC (annual financial activity quantity per capita) 15000 yuan, converted to node E; service area b has 15000 people, 34.56 square kilometers, AFAQPC 20000 yuan, converted to node B; service area c has 2000 people, 33.48 square kilometers, AFAQPC 14000 yuan, converted to node D; service area d population 16000 people, 69.12 square kilometers, AFAQPC 19000 yuan, converted to node F; service area e has 2800 people, 33.48 square kilometers, AFAQPC 17000 yuan, converted to node G; service area f has 2600 people, 33.28 square kilometers, AFAQPC 25000 yuan, converted to node J; service area g has 2800 people, 33.48 square kilometers, AFAQPC 18000 yuan, converted to node R; service area h has 2500 people, 33.28 square kilometers, AFAQPC 21000 yuan, also converted to node P; service area i has 2000 people, 21.12 square kilometers, AFAQPC 16000 yuan, converted to node O; service area j has 2800 people, 33.24 square kilometers, AFAQPC 15000 yuan, converted to node Q; the service area k has 3800 people, 61.14 square kilometers, AFAQPC 13000 yuan, converted to node V. The first number in brackets after each node is the population. The second is its area. The third is AFAQPC (the annual per capita financial activity quantity). It should be noted that according to the rules, the service area g and h are converted to the same node P, the number after the node P is the sum of the corresponding number for the two service areas, wherein the first two numbers were obtained by simple addition of the corresponding numbers of two service area, or the third number is gotten (rounded to one hundred yuan) by the sum of total annual financial activity of two service areas divided by the sum of the population of two service areas. The basic topological graph transformed is as shown in Fig. 2.

By dividing the basic topology, nodes is divided into different partitions, thus the financial services area division is achieved. This division will make the service numbers, service areas, or financial activity of the partition achieve the intended goals. For populous or financial activities large areas, division will make a large population and small area in the partitions; for sparsely-populated areas, division will make area slightly larger but not too big. Thus, in general, division will make financial services object to get the most satisfaction, at the same time, it will make each service terminal of each financial institutions to play maximum efficiency.

**B. Data structure of service area and topology graph**

The original service area data graph should include population, service area, and annual financial activity quantity per capita (AFAQPC) etc. The data file format of the service area is as the following:

The 1st line: d, 3.

The line of i+1 (i = 1 to d): n,s,F,x,L1-R1,……,Lx-Rx.

Each parameter represents the means as following:
d: service area number, 3: service area property number
The No. i+1 line has the No. i service area
n: population, s: area, F: annual financial activity quantity per capita (AFAQPC), x: the Number of the Gate in service area, L1-R1: gate 1’s left node and right crossroads of its street.

It makes the original graph transformed into the basic topological graph composed of nodes and edges. Here, the information of node includes population, service area, the annual per capita financial activity and so on.

Through calculating the service area data file, road intersections are abstract to vertexes, roads are abstract to edges, and topology graph G(V, E) can be got. The vertexes should have the 3 key properties that have been mentioned.

The file about to be partitioned has the format as the following:

The 1st line: v, 3.

The line of i+1 (i = 1 to v): w1, w2, w3; v1, e1, e2, ……, vP, eP.

Each parameter represents the means as following:
v: vertexes number; 3: vertex property number
The No. i+1 line has the No. i vertex weight and its edges information.
After compression, algorithm and centralized algorithm two categories. A seen as a static graph partitioning. Static graph partitioning the rural financial information services area division can be constructed. The BLP algorithm, etc. The hash algorithm and the BHP algorithm only considers the load balance, without considering the internal structure of graph, it's easy to appear the structure fracture phenomenon. Static Mizan algorithm categorizing graph processing respectively, cost is larger. The BLP algorithm can ensure the partition balance, and can guarantee the connectivity of the internal structure of the partition, but the time complexity is bigger [25]. The centralized algorithm mainly has KL/FM algorithm, spectrum dichotomy and multilevel k way partitioning algorithm. KL/FM method is the associated nodes together when making graph partitioning, without concerns the geometric distance between the nodes therefore it isn't conducive to partitioning need according to geographic information. Spectrum-dichotomy is to solve by the Laplace matrix of a graph calculating the corresponding eigenvectors, having a high partitioning quality, however, the computational complexity is also high. The multilevel k way partitioning algorithm is easy to realize large-scale graph partitioning, having faster speed and good partitioning quality [26]. It is able to ensure a balance of the division, but also to ensure a interior connectivity of the partition. In this paper, multilevel k way partitioning algorithm is used.

Main idea of multilevel k way graph partitioning algorithm is that first, the original graph is coarsened into a smaller graph, then this smaller graph is partitioned into k parts, finally, by improved refinement get a partitioning of the original graph. The idea of the algorithm is clear, formal is normative, software is forming. So based on this algorithm, a theoretical framework for graph partition algorithm of the rural financial information service area is constructed.

The operation process of graph partitioning method of rural financial information area includes three parts[17]: first is mapping transformations of the graph, constraint condition and the determination of the optimization goal; second is graph partitioning, including three stages which are the coarsen compression, the initial partitioning and refinement reduction.

A. Coarsening

Coarsening—is compressing the original graph into coarsen graph with less nodes. After compression, calculation can be faster and more efficient. According to the original graph \( G_0 = (V_0, E_0) \) to construct a series of the coarser graph \( G_i = (V_i, E_i) \), it was asked to \( |V_i| < |V_{i-1}| \) requirements. Can use the matching method for edge compression. Using light vertex matching algorithm to compress more suitable for partitioning a financial services area.

Two problems should be noticed: 1. the graph that has been coarsened will reduce less and less in the continued process. 2. Partition maybe unbalanced. But all kinds of partition and coarsen method work together will produce an excellent result.

B. Initial Partition

The initial partitioning is making the coarsest graph has been compressed is divided into k parts. the coarsest graph \( G_m = (V_m, E_m) \) is partitioned into k parts. The weight of vertices in each partition roughly equals to \( |V_m|/k, |V_0| \) is the sum weight of the original graph vertexes. The size of coarsest graph \( G_m \) is very smaller, so that the initial partitioning process will take less time.

Characteristics of the road connection also have some characteristics of the network connection, trunk road is the mesh structure, the country road presents like tree branches, so the network topology graph partitioning algorithm can be learn from. Using selective graph growing partitioning algorithm and depth-first and breadth-backing-off algorithm to make initial partitioning of coarsest graph \( G_m \) has a good division results.

C. Refinement

Refinement reduction is that the graph obtained after initial is gradually reverted to the original graph layer by layer, according to the inverse direction of coarsening; in the process of reduction, by adjusting nodes between the partitions, improving layer by layer, that the refined partitioning graph is obtained. K way refinement is very complex. a vertex can move from a part to the other k-1 parts, not like the bisection refinement that only 1 part. Thus it makes many chances to be optimized. The algorithm is using k(k-1) arrays to get the optimized movement. Each array is served for one movement. In each step the biggest gain is found in the k(k-1) arrays, and move the corresponding vertex and update all the arrays.

When a vertex is moved from one part to another part, both edge cutting should be reduced and the unbalanced state should be avoid. A vertex can be moved must satisfied the following formula:

\[
W_v(b) + W_v(c) \leq W_{max}
\]

\[
W_v(a) - W_v(c) \geq W_{min}
\]

In the formula, \( W_v \) is a vector which has k elements, \( W_v(a) \) is the weight of part a in graph \( G_0 \), \( W_{max}=0.9|V_0|/k \), \( W_{min}=c|V_0|/k \), \( W_v(c) \) is the weight of vertex v.

The k sets which have vertexes are the k partitions for the whole region, each service point is in charge of the
corresponding set, thus the whole financial service area is partitioned.

V. Evaluation of the Rural Financial Information Service Area Division

A. Evaluation Principle of the Rural Financial Information Service Area Division

In order to correctly evaluate the division quality, thus to compare the results of different divisions, it is needed to construct a evaluation index of the rural financial information service area division. The correct division principle is the basis of constructing the evaluation index of information service area division. The division principle of rural financial information service area include: profit maximization principle; cost minimization principle; supply and demand balance principle; spatial accessibility fairness principle; comprehensive coordination principle. The reason for the study of the comprehensive coordination principle, because in the process of the actual division of rural financial information service area, Some principles are sometimes contradictory, complete achievement is impossible, Therefore, must take into account all aspects, comprehensive coordination.

Designing evaluation indexes of rural financial information service area division, should be based on the weight of influence factors and principles of graph partitioning, from all angles of profit and cost, supply and demand, such as convenience and fairness, overall consideration, to investigate the role and relationship of each evaluation index, to determine the evaluation effect of each evaluation parameter for the graph partition quality, to build the construction of overall evaluation indexes.

General evaluation index should be able to reflect the division effect from the comprehensive quantification of different angles, should reflect the influence of division results of rural financial information service area to the service quality. The specific principles of constructing the evaluation index are:

First, the short distance principle: should make the service object at the distance cost as small as possible, that is, the distance of the service objects to reach service outlets as small as possible.

Second, the average time principle: For service objects of different partitions, the waiting time in service outlets should not differ too much. That is, the total number of each partition service should not be too much difference.

Third, the high availability principle: utilization of the service capacity of financial service agency should be as high as possible. That is, the total financial activity volume of a partition cannot be too small. To this end, the total financial activity volume of each partition should not be too much difference.

Fourth, the overall consideration principle: with multiple evaluation indexes, there may be a contradiction between the indices, therefore, it should take into account the above indices, through the mathematical method of fuzzy optimization, given weight coefficients of different indexes, integrated harmonious, objective function for the overall evaluation is given.

B. Evaluation Method of the Rural Financial Information Service Area Division

In order to implement the evaluation principle of rural financial information service area division, it is need to build a proper evaluation function, and provide operational evaluation method.

First, in order to implement the principles of the short distance, it is need to build an index parameter positively correlated with the total distance of all service objects to the service outlet. To this end, on the basis of the product of each partition service radius (r) and the service number (n) are calculated (the product of number and radius for short), it makes to calculate the sum of the product of number and radius of all partitions (\( \sum_n r_i p_i \)). Making

\[
J = \sum r_i p_i
\]

(3)

The division should make J value as small as possible.

in the above formulas, service radius r can be calculated based on area of the partition. The forms of partition may be not a circle, but it can be transformed into a circle with equal area, the radius of circle is the service radius r of the partition. Making the area of a partition is s, that is the area s of the circle transformed. According to the formula of the circle area \( s = \pi r^2 \), we can get

\[
r = \frac{\sqrt{s}}{\sqrt{\pi}} = 0.565 \sqrt{s}
\]

(4)

Secondly, in order to implement the average time principle, it is required that the total number of the service object of a partition cannot be too large. To this end, it is making the ratio of total number \( n_{\text{max}} \) of the service object of the maximum partition to the average one \( n \) be called the number of difference ratio \( p \), and \( p \) is not greater than the given value \( \lambda_p \). That is,

\[
\frac{n_{\text{max}}}{n} = p \leq \lambda_p
\]

(5)

Thirdly, in order to implement the high availability principle, it is should make each partition try not to waste the service ability, the total financial activity volume of minimum partition be not too small.

To this end, it is making the ratio of total financial activity volume \( F_{\text{min}} \) to the minimum partition to the average one \( F \) to be called the activity of difference ratio \( q \), and \( q \) is not less-than the given value \( \lambda_q \). That is,

\[
\frac{F_{\text{min}}}{F} = q \geq \lambda_q
\]

(6)

Finally, in order to implement the overall consideration principle, an overall evaluation objective function \( E \) which contains the parameter \( J \), \( p \), \( q \) can be constructed, and \( J \), \( p \), \( q \) with different weight coefficients in the formulas, the purpose of doing so is to coordinate degree of the role of each parameters in partitioning. The overall evaluation objective function \( E \) can have a variety of different forms of construction. If these three parameters are treated equally, it can make

\[
E = \frac{J \cdot p}{q}
\]

(7)

The division should make \( E \) value as small as possible.
VI. CONCLUSION
In this paper, by means of one-to-one mapping, the factors of the activity place (area) for rural financial information service object, including population, area and financial activity quantity are converted according to certain rules, and the basic topology of the service area is obtained. On the basis of, using multilevel k-way partitioning algorithm, the nodes are divided into different partitions, and the division of the rural financial information service area is implemented. The method both to ensure a balanced partition of the service area, but also to ensure the connectivity of partition interior; for the optimization of rural financial information network layout, provides a scientific decision-making tools; applications of graph partitioning method is extended.

ACKNOWLEDGMENT
Fund projects: National natural science foundation of China (71340022), science projects of China national ministry of information industry (01XK230009), major science and technology project of education department of Heilongjiang province (10511z015).

REFERENCES