A clique-based and degree-based clustering algorithm for expressway network simplification problem

ZIPENG ZHANG
School of Management science and Engineering,
Shandong Normal University,
Jinan, China,
HONGGUO WANG
School of information science and Engineering,
Shandong Normal University,
Jinan, China
zhangzipeng20@126.com

Abstract:-In order to meet the complex topological structure and huge distance and site number information of the expressway network, this paper proposed a new degree-based maximal clique mining algorithm which based on the principle from top to down to reasonably simplify the expressway network information. At the same time a search tree model which containing a series of pruning strategies and dictionary ordering strategies is derived to describe the mining process directly and improve the efficiency of searching and clustering the expressway network. Finally in this paper, a simplified algorithm of the expressway network model is designed, and the new algorithm shows better results based on the example databases of expressway network in Shandong province.

Keywords:-maximal clique mining; degree-based principle; distributed search tree model; expressway network simplification

1 Introduction

The construction of expressway can reflect the level of the traffic and the economic development in a country or area. China's expressway mileage has exceeded 96,000 km to the end of 2012 which is second only to the United States. In the context of the rapid development of transportation industry, the expressway transportation become an important part of the transportation industry due to its features of high driving speed, large carrying capacity, low traffic accident rate. However, with the continuous expansion of our country highway road mileage, it also has a phenomenon of much densely abuse and low efficiency of monitoring and management. Thus, to simplify the information and the complex topological structure of the expressway network is an effective way to solve the above problems.

But the current scholars for expressway research mostly focused on the aspects of the network reliability, the hierarchical division, the evolutionary pattern, and the capacity of the expressway network [1-2], so rarely involves the highway road network simplification algorithms. Inspired by the clustering thought which based on data reduction algorithm, this paper cited a maximal clique clustering thought which in the field of group mining to simplify the highway road network.

Maximal clique mining problem is a branch of graph mining which can be earliest traced back to a theoretical article which was published by the British scientists Luce & Perry in 1973[3.] Since it was first put forward not until the recent years maximal clique mining problem has always been attracting much attention from researchers in the group mining community. Because the objects of applications by the mathematic and natural model in the group theory are different, various maximal clique mining algorithms have been put forward by experts in various field in recent years[4-6], such as the social network problems, the biological information network, the intelligent transportation network, the wireless sensor network problems, the DNA interactive network problems and so on.

At present, scientists have put forward some effective maximal clique algorithm, such as the local search algorithm-MCODE[7] which is based on density in a biological network to find out the complex organism enginery of protein effectively and used in the field of social networks. K-medoids algorithm[8] for mining network connection structure based on node clustering is proposed to enumerate the community associated with the association. Meeting the problem that many networks display community structure—groups of vertices within which connections are dense but between which they are sparser and sensitive, Newman put up a fast algorithm[9] to detect the maximal structure. Some integrating graph clustering method are developed to capture protein modules/protein complexes by multiple network features in different algorithms and also prove the rationale to identify protein complexes and functional modules by detecting clique from a high coverage PPI network[10-11].

Although these algorithms have their own advantages, there is also one common shortcoming that the overlap and interaction of maximal clique structure cannot be recognized in the real network data. And the overlap and nest of the maximal clique structure are widespread in the real network.

The maximal clique enumeration algorithm discussed above is only for the general sense of the graph structure, whose meaning is to ignore structure itself to design a general algorithm of the given graph. And the effects of evaluation algorithm by data sets are simple random generation group based on artificial map.

The paper put the reality expressway traffic network topological structure as a starting point, and the maximal mining algorithm is applied to the expressway network to simplify the road network and reduce the computational complexity of the objective. An empirical study of mining and analysis of the maximum clique which the expressway contains is conducted.

2 Definition of the Problem

Before maximum clique algorithm in section 3 is presented, some definitions in the algorithm are to be introduced, which involve theorem, rules, conditions and description in detail. Relevant performance metrics and assumptions used in our work are given as follows:

2.1 Definition of the expressway network in the graph-theory

The expressway network is a large graph data model which contains a large number of entries and exits of stations, service stations and overpass.

When the graph-theoretic problems considered, the expressway network is defined as a tuple G = (V, E, W), where V denotes the set of entries or exits of the stations and the service stations in the expressway network, and E is a set of disorder two-tuples composed of the elements in V, which represents a set of existing expressway connected between two stations. W is the weight of edge E, which represents the distance between the station. E[i][j]=1 represents that V_i is connected to a V_j, else E[i][j]=0 says that Vi and V_j is not connected. And the definition of E[i][i]= ∞ means there is no loop in the same station. W[i][j]=2 points out V_i connected with V_j and the distance between them is 2, W[i][j]=0 saying that between point V_i and V_j, they are connected but the distance of them is 0 (here i, j $\in$ V(G)). D_i represents the degree of the node i, that is the number of nodes which is adjacent with the node V_i.

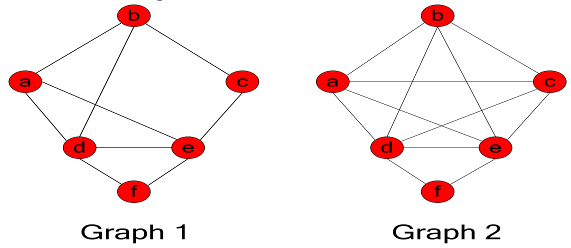

Figure1. Example for graph database

	a	b	c	d	e	f
a	∞	1	0	1	1	0
b	1	∞	1	1	0	0
c	0	1	∞	0	1	0
d	1	1	0	∞	1	1
e	1	0	1	1	∞	1
f	0	0	0	1	1	∞

Adjacency matrix for Graph1

	a	b	c	d	e	f
a	∞	1	1	1	1	0
b	1	∞	1	1	1	0
c	1	1	∞	1	1	0
d	1	1	1	∞	1	1
e	1	1	1	1	∞	1
f	0	0	0	1	1	∞

Adjacency matrix for Graph2

Figure2. Adjacency matrix representation of G_1 and G_2

Figure 1 is a graph theory representation of traffic network information data. The two graph in Figure1 are all composed of six points (a, b, c, d, e, f), but the topology displayed by the two graphs is not the same, because the connection between any two node are different. Figure 2 is an adjacent matrix representation of Figure 1. If node a and b is connected, then E[a][b]=1, and if node b and e did not connect, then E[b][e]=0. According to the graph theory there are no loop between the same point, so E[a][a]= ∞. Besides we can see the node d and e are with the biggest degree in graph1, whose number of the adjacent points is 4 so D_d=4, D_e=4.

2.2 Definition of the clique problem in the expressway network model

In the undirected graph of the expressway network G, if the sub-group Gs= (v, e) meets the condition that v $\in$ V(G), e $\in$ E(G), then G_s is called the sub-graph of G. If $C \in V(G)$, $a, b \in C$, $E[a][b]=1$, and

$E_{ab} \in E(G)$, then C is called complete sub-graph of graph G, and it is also called clique. A maximal clique is a clique that can not be extended by adding one more adjacent vertex, that is, a clique is not a subset of any other cliques. For example, in Group2 of Figure 1, the maximal clique is C{a, b, c, d, e} and C{d, e, f}.

The structure of overpass in the expressway network can be divided into T-shaped overpass and cross overpass. On the basis of connectivity between point and point of expressway network, the topology structure of the overpass can be turned into the data in the graph theory model. This model is also the starting point of maximal clique mining in expressway network. Figure 3 and Figure 4 describe the conversion process mentioned above.

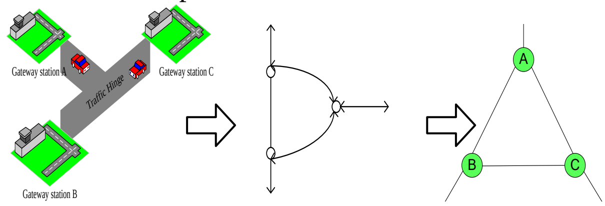

Figure3. Clique model of T-shaped overpass

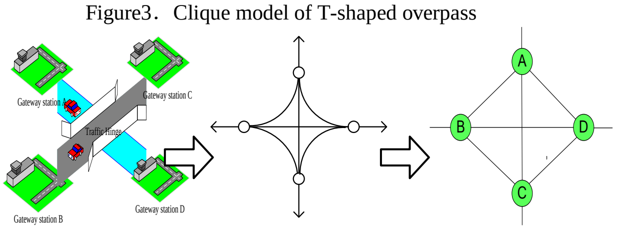

Figure4. Clique model of cross overpass

3 Realization of the Algorithm

3.1 Basic definition of the degree-based algorithm

In this section, we will discuss the algorithm designed for the special topology of the expressway network in detail, that is, a maximal clique mining algorithm which is a top-down algorithm, beginning from the vertex with maximum degree and then mining recursively all figure from the undirected graph, that is the degree method. The following Table 1 is a list of symbol and definition of the algorithm.

Table1 Symbol and definition in the algorithm

Symbol	Definition
G	undirected graph
Gs	sub-graph
i	node number
V_i	a node
$V(G)$	the set of nodes in G
$E(G)$	the set of edges in G
$N(V_i)$	$N(V_i)=\{V_j \mid E_{ij} \in E(G),\ V_i,\ V_j \in V(G)$
$N(V_i)'$	$N(V_i)' = N(V_i) - V(C_s)$
Di	$D_i = \mid N(V_i) \mid$, the degree of the node V_i
$D(G)$	the set of nodes with degree=D_i in G
$V(G)'$	a set of node V_i by descending Di
$N_L(V_i)$	$N_L(V_i)=\{V_j \mid V_j \in N(V_i), i<j\}, N_L(V_i) \in N(V_i)$
$\Delta(V_i, V_j)$	the set of nodes which can form a triangle with node V_i, V_j
C（G）	the set of all maximal clique in G
P_i	number of occurrences of the node i in C (G)
$E_{i,j}$	the edge connected by node i and j
$C_i\{i,j,k\ldots\ldots\}$	presentation of the maximal clique
$E(U)$	number of edges in maximal clique
$\sum W(E_{i,j})$	weight of edge $E_{i,j}$
$W(U)$	average weight
s	the size of the maximal clique
$V(C)$	the set of the nodes which belong to the maximal clique
$\mid V(Gc) \mid$	the number of nodes in the candidate clique

3.2 Basic idea of the degree-based algorithm

The degree-based algorithm is described as follows: firstly traverse the degree D_i of all points in the set of node V (G) using the depth first search method (DFS)[12], then select the node Vi according to the descending order of the degree, and then select the node V_j from the set N (V_i). Here we should find all possible nodes that can constitute a triangle with V_i and V_j, that is, there is a node $V_k \in \Delta(V_i, V_j)$, which satisfies the condition of $E[k][i]=1 \wedge E[k][j]=1$. Then we make V_j and V_k as the starting point for further recursively finding all possible nodes $V_m \in \Delta(V_i, V_j)$ $m \neq k$ satisfying the condition of $V_m \in \Delta(V_j, V_k)$, and $E[m][j]=1 \wedge E[m][k]=1$ from the collection E[m][j]=1 and E[m][k]=1. So the node V_i, V_j, V_k, V_m constitute a clique with the size of four. Because of $V_m \in \Delta(V_i, V_j)$ and $V_m \in \Delta(V_j, V_k)$, the whole process will go ahead recursively in turn until no any triangle structures is no longer found. Finally we will get a search tree that sees the node V_i as the root, in which each internal node respectively corresponds to a clique structure, and the leaf node represents that the node V_i can be contained by several cliques at all possible conditions. Of course, an undirected graph can contain many roots of search tree.

3.3 Lemmas related to the algorithm

- Theorem1: For any candidate, sub-clique structure $C_s\{i, j, k...... \}$ in the undirected graph G, if $\forall V_i \in C_s$ and only if there is no node $u \in N(V_i) - V(C_s)$ meet the condition $V(C_s) \subset N(u)$, then the structure of the candidate sub-clique C_s is the maximal clique.
- Theorem 2: In all candidate sub-clique structure $C_s\{i, j, k...... \}$, for $\forall V_i \in C_s$,always meets $V(C_s) \subset N(u)$.
- Theorem 3: Maximal clique mining problem can be divided into independent subtasks further, and each subtask is to mine corresponding maximal clique in search trees which regard each node as the root.
- Pruning rule 1: In the candidate sub-clique structure $C_s\{i, j, k...... \}$, we will select the node $V_i \in V(C_s)$ with the minimum degree D_{min} to use the decision rule of theorem 1 to determine whether the sub-clique C_s is the maximal clique.
- Pruning rule 2: In the process of extending the candidate sub-clique structure $C_s\{i, j, k...... \}$, for the node $\forall u \in N(V_i) - V(C_s)$, we always choose the node with $D_i \geq |V(C_s)|$ to expand the candidate sub-clique structure further.
- Pruning rule 3: In the process of mining maximal clique, the nodes in set V (C) contained by maximal clique C (G) cannot be used as the roots of the search trees, which is $V(G)' = V(G)' - C(G)$.
- Pruning rule 4: In order to avoid finding the maximum clique in the set $N(V_i)'$, if $|N(V_i)'| \geq 2$, we will go on mining maximum clique.
- Dictionary order 1: The node whose degree is higher than 2 in the set V (G)′ is in descending order. This dictionary order sets the degree of node in descending order, which can benefit for matching the top-down method of the maximal clique mining.
- Dictionary order 2: The nodes in adjacency node set N (V_i) are ordered by ascend according to their node label. Every time, for a given node V_i, the node with minimum label number is selected from the set N (V_i), so the entire search space are organized according to the dictionary order, which can effectively avoid searching the same candidate clique.

3.4 Process of the degree –based algorithm

- Step 1: With the depth first search scanning method, we first calculate the degree D_i of every node V_i in the database. According to the dictionary order 1, after the degree of the nodes in descending order, we select nodes who have the degree D_i>2 to form a new node set V (G)′. If $D_i=D_j=D_p=D_q$......, then the nodes are in descending order according to the node label.
- Step 2: Find out the adjacency node set N (V_i) of the node V_i with a maximum degree.
- Step 3: According to the dictionary order 2 remove minimum number in the adjacency node set N (V_i) of the node V_i, and select the node V_j, which meets the condition that $E[k][i]=1 \wedge E[k][j]=1$ from the rest node, and then select the node V_k whose node label is minimum and meets the condition that $E[k][i]=1 \wedge E[k][j]=1$, and then V_k can be added to the maximal clique C{i, j, k}. Next continue selecting the minimum nodes label from the remaining nodes who meet the condition that $E[m][j]=1 \wedge E[m][k]=1$, and then can be added to the maximal clique c{i, j, k, m}. Followed by recursion, we will finish until cannot find the node meets the conditions, and finally we will record the maximal clique C{i, j, k, m...... }.
- Step 4: According to the pruning rule 3, for nodes in the adjacency node set N (V_i) of the node V_i, we remove the nodes contained in the maximal clique to obtain the remaining node set N (V_i) '. Then return to step 3 until the rest node sets can no longer form any maximal clique. Now we can conclude that all cliques which consist of node V_i have been found.
- Step 5: Remove V_i and the set N (V_i) in V (G) ', and then select another V_i according to degree D_i in descending order again.
- Step 6: Repeat step 2, step 3 and step 4 until the node set V (G) 'is Φ.
- Step 7: The maximal clique set mined is sequentially numbered, then end.

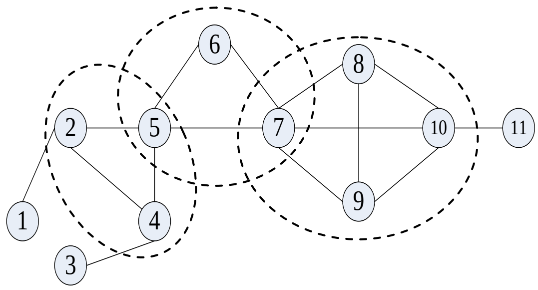

Figure5. Example of an algorithm

Figure5 is an example of the algorithm mentioned above. First we can get the degree D_i (for example D_1=1, D_2=3) of each node after traversing graph G, and then according to the degree, a new node set V (G) '(V_7, V_5, V_{10}, V_2, V_4, V_8, V_9) is formed. We start with choosing the node V_7 with the biggest degree in V (G)', finding N (V_7) = { V_5, V_6, V_8, V_9, V_{10}}, choosing V_5 in N (V_7), then select the node V_6 which satisfies the condition that E[i][5]=1 ∧ E[i][7]=1 in the rest node set{V_6, V_8, V_9, V_{10}} of N (V_7). And then select the node which satisfies the condition that E[i][6]=1 ∧ E[i][7]=1 from the remaining node set {V_8, V_9, V_{10}}, and the results is Φ, then the maximal clique found is U_1 {V_5, V_6, V_7}. In the next turn, we select the node V_8 with the lowest degree from the rest adjacency node set N (V_i) '= {V_8, V_9, V_{10}}, then select node from set {V_9, V_{10}}which can meet the condition that E[i][7]=1,E[i][8]=1 from the remaining nodes in N (V_i)'. We select node V_9 with the minimum number into group U (V_7, V_8, V_9), and finally we can find group U2 (V_7, V_8, V_9, V_{10}), because of the node V_{10} satisfies the condition that E[i][8]=1 ∧ E[i][9]=1. Then removing {V_5, V_7, V_{10}, V_8, V_9}from the set of node V (G) '(V_7, V_5, V_{10}, V_2, V_4, V_8, V_9) to get a new node set V (G)' (V_2, V_4), first select the node V_2 and then get the adjacent node set N (V_2) = {V_1, V_4, V_5}, the node V_1 is selected, and then select the node which can meet the condition that E[i][1]=1 ∧ E[i][2]=1 from the remaining nodes. Because the results is Φ, next the node V_4 is selected, then select node from the node set {V5} which can meet the condition that E[i][2]=1 ∧ E[i][4]=1, and finally maximal clique U_3 (V_2, V_4, V_5) is found. So at the end the maximal clique U_1 {V_5, V_6, V_7}, U_2 (V_7, V_8, V_9, V_{10}), U_3 (V_2, V_4, V_5) are found out.

On the basis of the idea and process of maximal clique mining algorithm we can also get the entire search tree space[13] considering corresponding nodes as the root, so it can effectively avoid repeating to search the same candidate clique. this paper proposes a parallel optimization method based on distributed system thought for the large scale search tree through the method that divide the set V(G)´ into equal size to use distributed system idea.

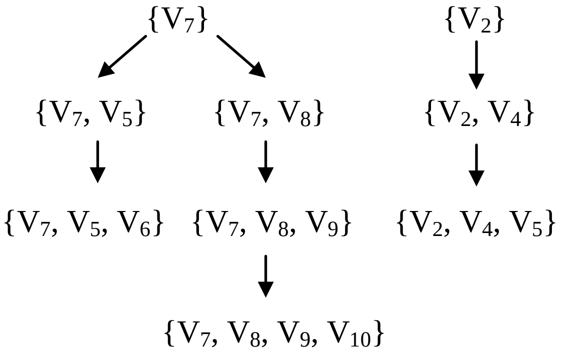

Figure6. Model of the search tree

4 Simplification of the Expressway Network

4.1 Idea of the simplification in expressway network

All discussion above which based on maximum clique mining problem is carried out under the premise of the undirected graph and do not be considered the weight in reality problem. For any node V_i in maximal clique C, the importance in the maximal clique can be defined as Formula (1)and Formula (2)[14]:

$$C_F(\mathrm{v_i}) = \frac{D_i}{(|V(C)|-1)} \tag{1}$$

$$C_F(G_c) = \sum_{vi \in V(Gc)} \frac{C_{F\max} - C_F(v_i)}{|V(G_c)|-2}, C_{F\max} = \max\{C_F(v_i)\} \tag{2}$$

By the topology and connectivity of the maximal clique and combined with the formula, it can be known that the importance of each node in maximal clique are the same, and the structure and properties of each node are also the same. As long as any one node is found out, we can find out any other nodes in the maximal clique. From Formula(2), the variable quantity $C_F(G_c)$ describes that the set of the candidate sub-graph has the trend of expansion, and if $C_F(G_c)$ is more, this trend will be stronger, and more likely to form a larger clique. The algorithm firstly constructs all possible triangles by merging all the clique of s=2 (i.e.: unilateral) and then further configured quadrilateral through the combination of triangle. Any triangle which can not be further constructed to be quadrilateral is judged to be maximal clique of s=3. The process executes constantly by using the iteration method until there no longer is a larger clique. The algorithm of Kros has performance deficiency similar to APriori. The first step of algorithm needs to traverse all edges of a given graph to construct a triangle, with complexity of O(M²)[15]. At the same time, in each stage of the algorithm, whenever we found a clique of sizes, we need to mark the size of the clique s-1 for non-maximal clique which is contained in it, so we also need to record all the clique of size s and s-1. It will take considerable storage if the size of the graph is very large.

In order to simplify the expressway network by the maximal clique mining algorithm, we should solve the problem of overlap and interaction of the maximal clique. Considering that we propose non-existent node V' to replace a true one, we see that the average weight of a maximal clique in (U) satisfies Formula (3)[16]:

$$W(U)=\frac{\sum_{i,j\in U} W(E_{i,j})}{E(U)}. \tag{3}$$

E (U) =s (s-1) /2

Based on the above, we combine the basic principle of maximal clique mining proposed in the third section with the idea of clustering, and regard the maximal clique as a clustering standard, and put the clique as a point to reduce the scale of data and the operation time of the network data to achieve the control of traffic network more efficiently and more conveniently. Figure 7 showing a demonstration model of simplified network algorithm and the nodes with same colour representing maximal clique structure. We simplify the 20 nodes of the original network to 10 nodes by our network optimization algorithm. Finally we found that the scale of the network is reduced by 50%, showing the efficiency and necessity of simplifying the network.

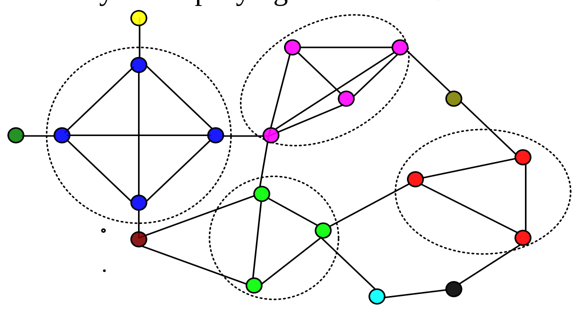

Figure7.Simplified model of the expressway network

4.2 Algorithm of the simplification in expressway network

Function: make the Simplified Group {G（V,E,W}

Input: the set of the maximal clique C(G)={C_i, C_j, C_k, C_m ……}

Output: make the simplified expressway network G'

Step1: traverse the set of the date C(G), record P_i as the number of occurrences of V_i ($V_i \in$ C(G))

Step2: if P_i=2, add V_i' as one of the node V_i, satisfied $V_i \in C_i$, $V_i' \in C_j$ and met the condition that E[i][j]=1,W[i][j]=0

If P_i=3, Add V_i' as one of the node V_i, V_i'' as one of the other node V_i, let $V_i \in C_i$, satisfied the condition that E[i][i']=1, E[i][i'']=1,E[i'][i'']=1,W[i][i']=0,W[i'][i'']=0,W[i][i'']=0

//select the node with $P_i \in$[2,3], update the maximal clique

Step3: calculate the edges of E（U）=s（s-1）/2 which in the maximal clique U_i（i, j, k, m……）

Step4: According to the following formula to calculate the average weight of the maximal cliques:

$$W(U)=\frac{\sum_{i,j\in U} W(E_{i,j})}{E(U)}$$

Step5: let V_i as U_i（i, j, k, m……）, re number of the set of V_i

Step6: receive the Simplified network G'

5 Experimental results and analysis

In order to display the effectiveness and performance of maximal clique mining algorithm which is based on the method of degree better, we use 443 site information and 1198 section information of expressway road network database in Shandong province. The experiment is conducted based on the following computer platform: Intel (R) Celeron (R) CPU E3300, 2.50GHz, dual core, 2G internal memory, the operating system of Windows XP Professional.

For the example like Table 2 which has 22 sites information around Ji'nan city, we can simplify it to Table 3 which has 5sites information.

Table2 Sites of original expressway network around Ji'nan city

Number	Station Name	Station Type	The Adjacent Site
1	Gang'gou	0	{2,336,335}
2	Pan'long	0	{1,3,335,336}
43	South Yu sity	0	{42,44,157,158}
44	Qi'jia	0	{43,45,157,339,158}
45	Ji'yang	0	{44,46}
46	Yao'qiang	0	{45,47,338,261,337}
47	Zhang'qiu	0	{46,48,337,261,338, }
157	Yu city	1	{156,158,43,44}
158	Qi'he	0	{157,159,339,316,43,265,44}
159	West Ji'nan	0	{158,160,265,266,316,335,339}
160	Gu'shan	0	{159,161,335,266}
261	Hua'shan	0	{262,338,46,337,47}
262	Ji'nan service area	1	{261,263}
263	Huai'yin	0	{262,264}
264	North Ji'nan	0	{263,265}
265	Tian'qiao	0	{264,266,316,339,159,158}
266	Chang'qing	0	{265,267,159,335,160}
316	Yan'cheng	0	{315,158,265,159,339}
335	South Ji'nan	0	{266,2,1,336,160,159}
336	East Ji'nan	0	{1,2,337,335}

337	Guo'dian	0	{336,338,261,46,47}
338	Ji'chang	0	{46,261,337,47}

Table3 Information of the maximal clique around Ji'nan city

Num	The Maximal Clique	The Information of the Clique
1	{1,2,336,335}	Gang'gou clique {Gang'gou, Pan'long, South Ji'nan, East Ji'nan}
2	{46,338,47,337,261}	Ji'chang,Xu'jia clique {Yao'qiang，Ji'chang，Zhang'qiu，Hua'shan，Guo'dian}
3	{157,43,158,339,44 }	Qi'he,Biao'baisi clique {Yu city，South Yu city，Qi'jia，He'dong，East Qi'he}
4	{265,339,158,316,159}	Huai'yin,Chi'zhuang clique {Tian'qiao ,He'dong，Qi'he，East Qi'he，Yan'cheng，West Ji'nan}
5	{160,266,335,159}	Yin'jialin clique {Chang'qing，South Ji'nan，West Ji'nan，Gu'shan}

It can be very efficient digging out the maximal clique in the Shandong expressway network with the degree-base algorithm. From Figure 8 and Figure 9, we can see that the complexity of the expressway and the station in the simplified expressway network is much smaller.

Figure8.Original expressway network in Shandong

Figure9.Simplified network in Shandong

It is a comparison of efficiency between original and simplified network in Figure 10. The first group is the comparison of time in depth first search method, and the second one is a comparison between the degree-based algorithm and enumeration method[17]. We can see that the degree of law and the introduction of a variety of pruning strategy lexicographic strategy greatly reduced the size of the algorithm running.

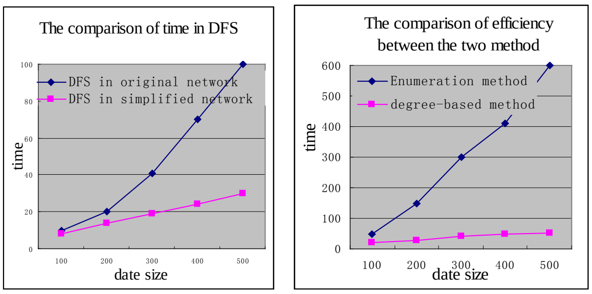

Figure10. Comparison of efficiency between original and simplify network

6 Conclusion

The paper proposes an effective expressway maximum clique mining algorithm based on degree to capture the maximal structures for the practical transportation network problem. The algorithm introduces the top-down search method innovatively, iteratively using pruning strategy, at the same time using parallel computing, so as to full-fill the processing of the data from the large-scale expressway network on the basis of reducing the complexity of time and space. This paper also presents a simplified model of expressway network to more fully show the superiority of the above algorithm. Through citing the databases of expressway network in Shandong province, this algorithm performs better than other available network clustering/mining methods in terms of the clustered members and the precision of obtaining Clustering clique. Using the maximal clique mining algorithm to simplify the national expressway network and monitoring expressway information in real time are the main research direction of this field. Importantly, this method can also achieve the same result by using the idea of distributed system to greatly reduce the computing time and complexity.

References:

[1]James Cheng, Yiping Ke, Ada Wai-Chee Fu, Jeffrey Xu, Linhong Zhu. *Finding Maximal Cliques in Massive Networks.[J]*. ACM Trans. Database Syst.,2011,36:

[2]David Eppstein, Darren Strash. List All Maximal Cliques in Large Sparse Real-World graphs[J].

coRR, 2011,abs/1103.0:
[3]Bron,C and J. Kerbosch: *finding All clique of an undirected graph*.ACM,1973.16(9):p.575-577.
[4]Kose,F.et al., *Visualizing plant metabolomic correlation net works using clique-metabolite matrices*.Bioinformatics 2001.17(12):p.1198-1208
[5]Lijun Chang , Jeffrey Xu Yu .*Fast Maximal Cliques Enumeration in Sparse Graphs Algorithm [J]*. 2013, Vol.66 (1), pp.173-186
[6]Xifeng Yan, Jiawei Han.gSpan: *Graph-based substructure pattern mining ［C］*.Proceedings of the 2002 IEEE International Conference on Data Mining. Washington, DC: IEEE Computer Society,2002: 721 － 724
[7]Eblen John D, Phillips Charles A, Rogers Gary L. *The Maximum Clique Enumeration Problem: Algorithm, Application and Implementations [J]*. BMC bioinformatics,2012,13 Supp 1:
[8]Weili Zhang, Xuhua Yang. *Dynamically Weighted Clique Evolution Model in Clique Networks.* Mathematical Problems in Engineering, 2013.Vol.
[9]Newman M E J. *Fast algorithm for detecting community structure in networks[J]*. Physical review E, 2004, 69(6): 066133.
[10]Chin C H, Chen S H, Chen C Y, et al. *Spotlight: Assembly of protein complexes by integrating g*raph clustering methods[J]. Gene, 2012.
[11]Wu M, Li X, Kwoh C K, et al. *A core-attachment based method to detect protein complexes in PPI networks[J]*. BMC bioinformatics, 2009, 10(1): 169.
[12]Eblen John D , Phillips Charles A , Rogers Gary L , Langston Michael. *The maximum clique enumeration problem: algorithms, applications, and implementations.* BMC bioinformatics, 2012
[13]Anne Berry, Romain Pogorelcnik. *A Simple Algorithm to Generate the Minimal Separators and The Maximal Cliques of a Chordal graph[J]*. Information Processing Letters,2011,11111:.
[14]Eblen, John D., et al. *The Maximum Clique Enumeration Problem: Algorithms, Applications, and Implementations.* BMC bioinformatics 13.Suppl 10 (2012): S5
[15]Li M, Wang J, Chen J. *A graph-theoretic method for mining overlapping functional modules in protein interaction networks[M]*//Bioinformatics Research and Applications. Springer Berlin Heidelberg, 2008: 208-219.
[16]Henry C J, Ramanna S. *Maximal Clique Enumeration in Finding Neighbourhoods [M]*. Transactions on Rough Sets XVI. Springer Berlin Heidelberg, 2013: 103-124.
[17]Boginski V, Butenko S, Shirokikh O, et al. *A network-based data mining approach to portfolio selection via weighted clique relaxations[J]*. Annals of Operations Research, 2013: 1-12.