Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SPATIAL TOPOLOGY ASSOCIATION MINING WITH CONSTRAIN CONDITION

[1]XU JIA-LIANG , [2]FANG GANG ,[3]YE XIAO-QIN

[1] College of Mathematics and Statistics, Chongqing Three Gorges University, China

[2]College of Computer Science and Engineering, Chongqing Three Gorges University, China

[3] Chongqing Three Gorges Medical University, China

E-mail: [1]ywing123@163.com, [2]fg_007@163.com, [3]yxq_leaf@163.com

ABSTRACT

Aiming at present mining algorithms can not extract spatial topology association with constrain condition, and traditional constrain association rules mining algorithms have repeated computing and superfluous candidate in the time of mining spatial topology association rules, an algorithm of spatial topology association mining with constrain condition is proposed, which is suitable for mining topology association with constrain condition in multi-spatial relation pattern. The algorithm uses topology binary to turn spatial relation transaction into integer, and uses value ascending of constrain condition to form candidate frequent item sets. To efficiently improve mining efficiency, the algorithm uses Boolean operation to compute support, and uses number attribute to reduce the number of scanned spatial transaction. The algorithm is used to mine spatial topology association rules with constrain condition, and this experiment indicates that the efficiency is fast and efficient.

Keywords: *Topology Association; Constrain Condition; Value Ascending; Topology Binary; Spatial Association Rules.*

1. INTRODUCTION

Since the expansion of the traditional association rules in the field of spatial data mining has three types of methods that layer covering method based on clustering and Mining method based on space affairs and Mining method based on non-transaction space was proposed. Mining method based on space affairs in a spatial database using spatial overlay, buffer analysis found that the predicate of space between the space goals and other excavated objects, space predicates Space Affairs in accordance with the mining target database, Boolean association rules Mine, Such as Documents[1, 2 and 3].These algorithms are able to tap a single space in the same spatial relationship mode association, but it can't effectively extract a space more than the relationship between a mode of the multi-layer spatial association, such as the typical spatial topology association; Documents[4, 5, 6 and 7], the proposed algorithm can effectively tap the space topology in the multi-relational schema associated. However, not be able to tap the space topology associated with the constraint condition; Separate algorithms proposed in Document[8] Although mining space topology associated with restrictive conditions, but because of double counting and redundant candidates, the efficiency of the algorithm needs to be improved.

2. WITH THE CONSTRAINTS OF SPATIAL TOPOLOGY ASSOCIATION MINING

Spatial topology association is the most complex space relationship model spatial association study which using a method based on space matters, so I do not consider the spatial topological between two objects are equal. This article only studied seven spatial topological relations in Table 1.

2.1 Space Affairs Conversion Method

Set up space for the target object Ot - non-target object m categories, denoted as ik (k = 1 ... m). According to the mining method based on space affairs and space affairs conversion method is described as follows:

Input: ①a target object Ot - m non-target object class (denoted by I = {i1, i2, ..., im});

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

② Binary topological relations of the corresponding table (Table 1) other (000) within said non-target object is not in the buffer of the target object;

③ A representation of topological relations of space affairs.

Output: an integer. Its specific conversion method is as follows:

①The topological relations in space affairs in accordance with Table 1 converted to binary form. That is, the topological relationships of target and non-target object three binary numbers to as a whole, called the topology of a binary number.

② the topology of a binary number constitute a binary number according to the position of non-target object class sort, and their median number of non-target object is three times of it.

③ converted to a binary number that consists of an integer, so that the affairs of each space will correspond to an integer deposited into the mining database.

Table 1 Space Topology Relationship Table

Serial number	Spatial topological relations	Predicate expressed form	Topological binary number
	O_t and i_k disjoint	disjoint(O_t , i_k)	
1	O_t and i_k touch	touch(O_t, i_k)	$(001)_2$
2	O_t and i_k overlapp	overlap(O_t , i_k)	$(010)_2$
3	O_t cover i_k	cover (O_t, i_k)	$(011)_2$
4	O_t contain i_k	contain (O_t , i_k)	$(100)_2$
5	O_t covered by i_k	coveredby (O_t , i_k)	$(101)_2$
6	O_t inside i_k	inside (O_t, i_k)	$(110)_2$
7			$(111)_2$

2.2 Space Topology Generation Of Candidate Frequent Method

The principle is in the mining of spatial topology association rules generate candidate frequent item: any material who contains the space constraints of a frequent (k +1) - itemsets, there are at least two have the same (k-1) contains the constraints frequent k-item subset of the specifically generated in two ways:

First, the major advantage of the connection function, that is, any two with space constraints (k-1) common k-itemsets connection, generate space constraints of the candidate (k +1) - itemsets;

The second is an expansion of generation with space constraints (k-1) - itemsets with the frequent 1 - itemsets.

Frequent itemsets length increments, but when frequent items contained in the number of spatial object's increases, the algorithm will produce a large number of redundant candidates and double counting. The efficiency of the algorithm will be affected.

The algorithm in mining association rules of spatial topology to build the candidate frequent item is: starting from the integer value representing the space constraints, incremental integer value, if it contains space constraints may be used as the candidate frequent item, and computing support; to increase the value to space affairs corresponding to the maximum value of the integer on the termination of generating candidate frequent item.

For example, Space Affairs set up the database to convert the integer, which corresponds to the largest integer for the 511, the space constraints corresponding to the integer 5, the candidate frequent item generation process :5,13,...,501,509.Which 6,7,8,9,10,11,12,...,502,503,504,505,506,507,508,510 and 511 do not contain constraints 2.2 of the nature, and inference can be deleted

By comparison: the algorithm generates candidate in a frequent way of the spatial topology and efficiently and effectively separate algorithm.

2.3 Method Of Calculation To Support The Number

Majority of spatial data mining algorithms calculate the support number, a candidate frequent item of data transactions between the support and the calculation method that used string matching, the set of legal and graphical superposition method to calculate the time support number, the majority of the algorithm need to scan the database of all non-repetition space Affairs, which will not be conducive to the improvement of algorithm efficiency. According to the nature of 2.2 shows, the algorithm in calculating the support number, first of all require a candidate frequent integer corresponding to the Space Affairs to conduct a "and" operation to meet the nature of the 2 "and" the result, then the length of the computation, the final order to determine the relationship between the two; calculation only needs to scan the database corresponds to an integer greater than or equal to the candidate frequent item integer space Affairs. For example, Space Affairs set up the database 20, the converted integer {234, 101, 98, 73, 63, 62, 59, 55, 49, 42, 31, 25, 22, 19, 17,15,13, 11, 7, 5}, if

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

space constraints corresponding to an integer 16, the candidate frequent item 18, only 14 digits of the number of computing support algorithm {234, 101, 98, 73, 63, 62, 59, 55, 49, 42, 31, 25, 22, 19} and calculation, and further calculate the length of only 7 digits {63, 62, 59, 55, 31, 22, 19}. However, this 5numbers {17, 15, 13, 11, 7, 5} do not need to be calculated, because the corresponding space affairs, cannot contain the 18 corresponding to the space Affairs.

2.4 Algorithm Of Mining Ideas

Database of symbols: TD: Storage Space Affairs;

D: for each space in the store database transaction, corresponding number and transaction number of repeats, contains two domains, the "value" and "count"Frequent item set's F: Storage mining process; NF: storage in the mining process, the non-frequent item sets;

Step1: According to knowledge in 2.1 section in the database TD Space Affairs to convert an integer, and descending stored in the D

Step2: According to the integer corresponding to the maximum value max and space constraints' IC, construct an integer interval [IC, max], the IC as a candidate frequent item C1, calculate the support count of frequent items deposited F, otherwise terminate the algorithm.

Step 3: increment the value of Ci, Ci to meet one of the above two conditions to continue to increase its value (① does not contain a space constraint; ② The NF neutron superset), otherwise as a candidate frequent item to support calculation, if the frequent items stored in F, otherwise deposited in NF.

Step4: Repeat Step3, until the Ci increments reach maximum max, the search for the frequent items concerning the termination.

Step5: the frequent items' F contains the constraints of the spatial topology association rules.

Step6: the above association rules restore the topological according to the method of 2.1 section.

Relations between spatial objects, the formation of spatial topology association rules contained constraints of the standard.

3. PERFORMANCE ANALYSIS OF ALGORITHM

3.1 Time Complexity Analysis

Space Affairs the number of n ($n \leq 23m$) set up the database does not repeat the space of non-target object number is m, the number of non-target space object space constraints contains the t ($t < m$), the corresponding integer p ($p \geq (23t-1) / 7$), the time complexity is expressed as:

$$T = \sum_{k=1}^{2^{3m}-(2^{3t}-1)/7} k \circ$$

Be seen: when t increases, the increase in the number of constraints contains a space object, contains the number of constraint's candidate frequent item will be reduced, the number of scanned Space Affairs will be reduced, so the calculation will reduce time complexity. T reduced the constraints contained in the number of spatial objects to reduce the constraints of the number of candidate frequent item and will increase the number of scanned Space Affairs, so the increase in the computational time complexity rise

3. 2 Space Complexity Analysis

Reduce the amount of space to store the algorithm, the algorithm using topological binary space affairs which is converted to an integer storage space complexity that can be expressed as O (q.nx M), M is the largest project set, q is a support and restraint conditions of the parameters.

3.3 Compare Of Experimental Results

The present study of the constraints of the spacial topology association rules mining algorithm does not use existing binding association rule mining algorithm Separate and the proposed algorithm STAMWC to experiment here, the more spatial topology association rules mining with constraints superiority.

The test data is as follows: total amount of data for 6142 and 4095 does not repeat the transaction number, the corresponding integer of 1-4095, the repeat number "1" and "2" alternating Space Affairs, that is, 4095 1, 4094 2 4093 1 4092 2 ... 4 in the space object. Space constraints, their spatial topological relations is expressed as a predicate form of touch (O, D), O is the target object D indicates the category of non-target came in the fourth position of spatial objects.

Algorithm used in the experimental environment: Intel (R) Celeron (R) M CPU 420@1.60 GHz, 1. 24G memory, the operating system for Windows XP Professional, two algorithms are implemented in Visual C # 2005, NET development platform Separate and STAMWC.

Separate STAMWC, algorithm execution time with the support of the change in the comparison is shown in Figure 2.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

From the experimental results of the comparison we can see that the spatial topology association rules mining with constraints, STAMWC higher than the Separate efficiency of the algorithm; Figure 2 STAMWC algorithm execution time is reduced with support and increase the frequent items that contain the space object to increase the number of run-time will also increase.

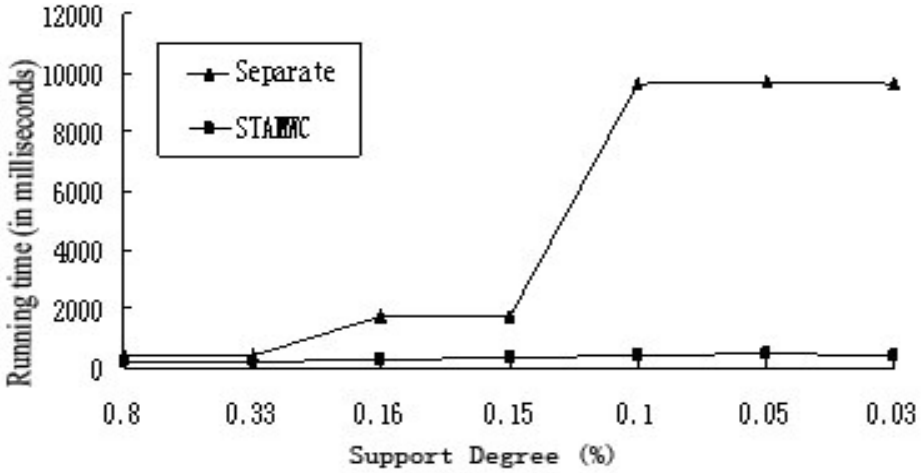

Figure 2 Compared to the situation of execution time with the project to support change

4. CONCLUSION

For the existing space, spatial topology association rules mining algorithm cannot extract with restrictive conditions, coupled with the tradition-bound algorithm for mining of association rules, there is double counting and redundant candidate, this paper proposes a kind of constraint spatial topology association mining algorithm STAMWC, the algorithm can effectively make spatial topology association rules mining meet the constraints. The experiments prove that the algorithm is more quick and efficient than existing mining algorithms.

REFERENCES

[1] Gang FANG, Zu-Kuan WEI, Qian YIN. Extraction of Spatial Association Rules Based on Binary Mining Algorithm in Mobile Computing[C]. IEEE Information Conference on Information and Automation. IEEE Press, China, 2008:1571-1575.

[2] Liu Yu-Lu, Fang Gang. A transaction complementary research and application of mining algorithms [J]. Computer Engineering and Applications, 2008,44(35):168-170.

[3] Fang Gang, Liu Yu-Lu. Mining algorithms in mobile intelligent systems based on binary space.[J]. Southwest University (Natural Science), 2009, 31(1): 95-99.

[4] Xiong Jiang, Fang Gang, Liu Yu-Lu. Two-way excavation of the spatial topology associated[J]. Computer Engineering and Applications, 2009,45(22):126-128.

[5] Tang Xiao-Bin, Fang Gang. For the topology of the horizontal space mining association rules algorithm [J]. Computer Engineering and Applications, 2010,46(1):109-111.

[6] Luo Ai-Ping. Spatial cross-layer association rule mining algorithm [J]. Southwest China Normal University (Natural Science), 2009, 34(4):1-5.

[7] Fang Gang, Wei Zu-Kuan. An effective algorithm for mining spatial topology association[J]. Computer Engineering and Design, 2010, 31(6): 1267-1270.

[8] Shao Feng-Jin,Yu Zhong Qing. Data mining principles and algorithms [M]. Beijing: Science Press，2009, 8.