ORGANIZING SPATIO-TEMPORAL DATA WITH MUTI-BASE-STATES AND MULTI-CLASSES DELTA FILES

ZHANG Baogang[a] ,ZHU Chongguang[b] ,WANG Runsheng[c]

[a] Geographic Information Centre, Surveying and Mapping Institute of Beijing, Beijing, China, bg_zhang@163.com [b] Institute of Remote Sensing Application, CAS, Beijing, China [c] China Iero Geophysical Survey & Remote Sensing Center for Land & Resource, Beijing, China

Key words: GIS ; multi-classes delta-files; multi-base states

ABSTRACT:

Abstract: "Base state with amendments" preserve data state of sometime and increment relative to base state. It reduces redundant data, but efficiency to recover historical data has changed lower. In order to improve efficiency to recover historical data, methods are proposed to construction multi-base states or multi-classes delta-files. This paper proposed a method to construction "multi-base states and multi-classes delta-files" synchronously.

1. INTRODUCTION

Various spatio-temporal data models have been proposed to store spatio-temporal data, such as snap shot model, base state with amendments, spatio-temporal composite model and so on. Base state with amendments only preserves base state in certain time and variance relate to base state. It records and maintains variance of spatial objects and topologic, and doesn't store all information of each state. Base state is the state of updating data state at last, the variance relate to base state is expressed by delta-file. Comparing with other spatio-temporal data models, this method is easy to realize. It is a problem for many scholars to solve to improve characters and make it fit to store spatial information changing with high frequency. Langran(1988), Zhang Z.(1996) and Cao Z. (2002) has proposed five different amendments forms based state.

Form a regards each historical state as a amendment relate to now. This method need to modify delta files each historical time to base state (Kennedy-Smith, 1986). It is very long-winded and difficult for people to receive when many times of variance happed. Form b gets rid of disadvantages of form a, Langran chose form b. However form b needs to read all delta files, when it is very long, computers have to read many delta files. Form b is not a perfect method also.

Form c and d are two kinds of improved methods proposed by Zhang Z. (1996). They overcome disadvantages in form a and b to some extent. Same problems existing in form a and b also exist in form c and d when spatial data is of old-line. Form e is a method of "Multi-base-states and single class delta files" (The following text call it MS) proposed by Cao Z. (2002). This method establishes a serious base states dynamically. Distance between to nearest base states is called base-state-distance. Base-state-distance can be determined upon users or systems according to concrete problem. With inserting and amending of historical states of objects, the system establishes base state dynamically according to frequency to use certain state. This mean new base state will be established automatically when base-state-distance is greater than a certain threshold with new variances adding. Recovering of historical states may spend a lot of time when users retrieve some states frequently especially to continuous spatio-temporal process of variances. It is solved

to construct base states dynamically to some extent (Cao Z.,2002).

Doing as Cao Z. (2002) abbreviate time of retrieving spatio-temporal data, however it increase storing space of base states. When spatial objects are long-line, the space is much larger. This paper is ready to organize spatio-temporal data with the method of "multi-base states and multi-classes delta-files". It can both save storage space and decrease time to recover historical spatio-temporal data.
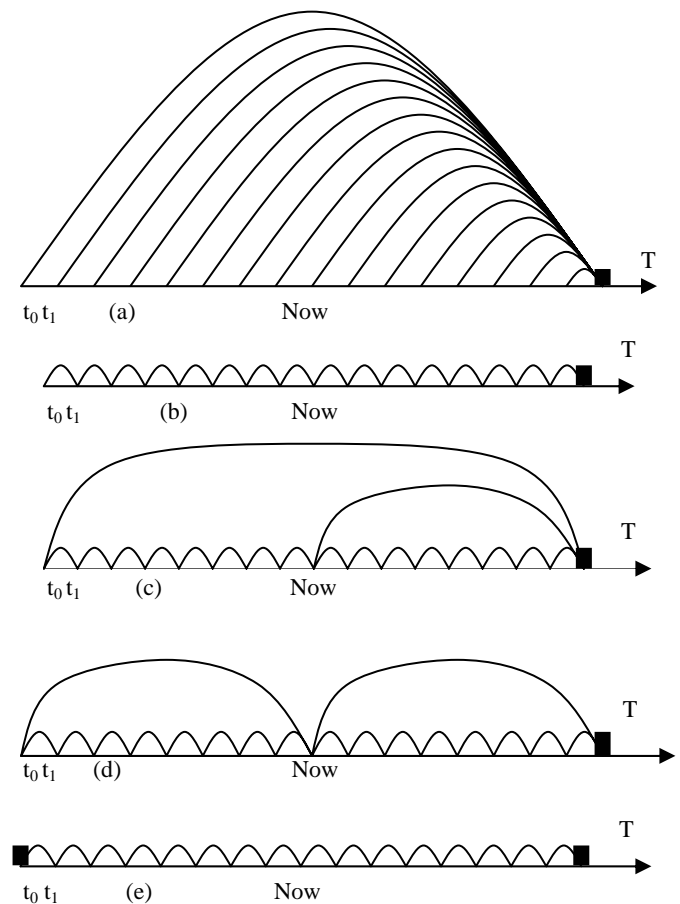

Figure 1. Five kind of base-state amendments

2. A METHOD OF "MUTI-BASE-STATES AND MULTI-CLASSES DELTA-FILES"

It can be find out that form c increases two classes delta files and form d increases one class delta files comparing with form b in Figure 1, form e increases multi-base-states comparing with form b in Figure 1. If form c and e or form d and e are combined together, It will form a result of "multi-base-states and multi-classes delta files" (The following text called it MM) or "multi-base-states and bi-classes delta files" (The following text called it BM) like form a and b in Fig2. Obviously, it will save more space to store delta files than to store base states.

Because of storing disadvantage of form a, b, c and d when spatial objects are old-line, it is useless to compare them with two forms in Figure2. If value of base-state-distance is threshold, base states and delta files of form e in Figure1 and form a and form b in Figure2 are compared in Table1. When threshold from 7 to 39, average numbers retrieving delta files with form e in Figure1 and form a or b in Figure2 are listed in Table2, Figure3 is the statistic graph of Table2. Table3 is the flowing route to reach each state. The direction of difference in delta-files is opposite to the direction of flowing route from arrowhead to arrow tail.
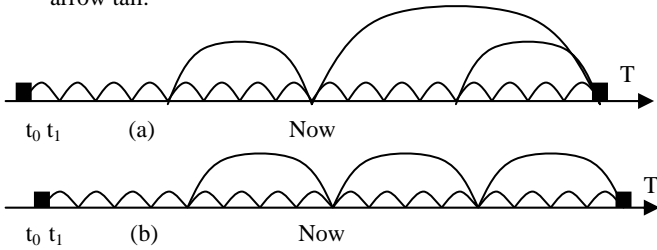

Figure2. "multi-base-states and multi-classes delta files"

Form of storage	Base-states number	Total delta files number	Longest delta files number
MS	2	threshold	[(threshold +1)/2]
MB	2	threshold+3	[(threshold +24)/8]
MM	2	threshold+3	[(threshold+16) /8]

Table1 Base-states number and delta files number of "multi-base-states and multi-classes" and "multi-base-states and single class delta-files"

Threshold	MS	MB	MM
7	2	1.500	1.250
11	3	2.000	1.583
15	4	2.313	1.750
19	5	2.600	2.150
23	6	2.833	2.417
27	7	3.143	2.679
31	8	3.375	2.938
35	9	3.667	3.194
39	10	3.900	3.450

Table 2 Average number to retrieve delta files with "multi-base-states and multi-classes" and "multi-base-states and single class delta-files"
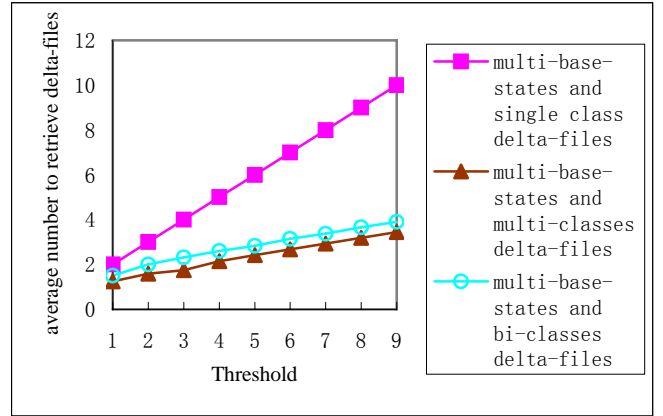

Figure 3. Relationship between average number to retrieve delta file and base-state-distance threshold of multi-base-states with amendments

	Figure1e	Figure2a	Figure2b
t_0	-	-	-
t_1	$t_0 \rightarrow t_1$	$t_0 \rightarrow t_1$	$t_0 \rightarrow t_1$
t_2	$t_0 \rightarrow t_1 \rightarrow t_2$	$t_0 \rightarrow t_1 \rightarrow t_2$	$t_0 \rightarrow t_1 \rightarrow t_2$
t_3	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3$	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3$	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3$
t_4	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4$	$t_{16} \rightarrow t_8 \rightarrow t_4$	$t_{16} \rightarrow t_{12} \rightarrow t_8 \rightarrow t_4$
t_5	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5$	$t_{16} \rightarrow t_8 \rightarrow t_4 \rightarrow t_5$	$t_{16} \rightarrow t_{12} \rightarrow t_8 \rightarrow t_4 \rightarrow t_5$
t_6	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5 \rightarrow t_6$	$t_{16} \rightarrow t_8 \rightarrow t_7 \rightarrow t_6$	$t_{16} \rightarrow t_{12} \rightarrow t_8 \rightarrow t_7 \rightarrow t_6$
t_7	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5 \rightarrow t_6 \rightarrow t_7$	$t_{16} \rightarrow t_8 \rightarrow t_7$	$t_{16} \rightarrow t_{12} \rightarrow t_8 \rightarrow t_7$
t_8	$t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5 \rightarrow t_6 \rightarrow t_7 \rightarrow t_8$	$t_{16} \rightarrow t_8$	$t_{16} \rightarrow t_{12} \rightarrow t_8$
t_9	$t_{16} \rightarrow t_{15} \rightarrow t_{14} \rightarrow t_{13} \rightarrow t_{12} \rightarrow t_{11} \rightarrow t_{10} \rightarrow t_9$	$t_{16} \rightarrow t_8 \rightarrow t_9$	$t_{16} \rightarrow t_{12} \rightarrow t_8 \rightarrow t_9$
t_{10}	$t_{16} \rightarrow t_{15} \rightarrow t_{14} \rightarrow t_{13} \rightarrow t_{12} \rightarrow t_{11} \rightarrow t_{10}$	$t_{16} \rightarrow t_{12} \rightarrow t_{11} \rightarrow t_{10}$	$t_{16} \rightarrow t_{12} \rightarrow t_{11} \rightarrow t_{10}$
t_{11}	$t_{16} \rightarrow t_{15} \rightarrow t_{14} \rightarrow t_{13} \rightarrow t_{12} \rightarrow t_{11}$	$t_{16} \rightarrow t_{12} \rightarrow t_{11}$	$t_{16} \rightarrow t_{12} \rightarrow t_{11}$
t_{12}	$t_{16} \rightarrow t_{15} \rightarrow t_{14} \rightarrow t_{13} \rightarrow t_{12}$	$t_{16} \rightarrow t_{12}$	$t_{16} \rightarrow t_{12}$
t_{13}	$t_{16} \rightarrow t_{15} \rightarrow t_{14} \rightarrow t_{13}$	$t_{16} \rightarrow t_{12} \rightarrow t_{13}$	$t_{16} \rightarrow t_{12} \rightarrow t_{13}$
t_{14}	$t_{16} \rightarrow t_{15} \rightarrow t_{14}$	$t_{16} \rightarrow t_{15} \rightarrow t_{14}$	$t_{16} \rightarrow t_{15} \rightarrow t_{14}$
t_{15}	$t_{16} \rightarrow t_{15}$	$t_{16} \rightarrow t_{15}$	$t_{16} \rightarrow t_{15}$
t_{16}	-	-	-

Table 3. States ransacking flow of spatio-temporal objects with the three methods in various time when threshold = 15

It can see from Table1 that "Multi-base-states and multi-classes delta-files", "Multi-base-states and bi-classes delta-files" comparing with "Multi-base-states and single class delta files", three delta files are increased without increasing the number of base-states, but the number of the longest delta files decline to a great extent. It indicates that recovering speed of spatio-temporal data enhance much more with increasing little storage space of spatio-temporal data. It means from Figure3 and Table2 : (1) the number of delta files is most using "Multi-base-states and single class delta files", and least using "Multi-base-states and multi-classes delta-files". (2) The more that the threshold value of base-state-distance is, the more that the delta files need to retrieving; the less that the threshold value of base-state-distance is, the less that the delta files need to retrieving. (3) The relation between average number need retrieve delta files

and threshold of base-state-distance is linear: slope with "Multi-base-states and single class delta files" is the largest. "Multi-base-states and bi-classes delta-files" as well as "Multi-base-states and single class delta files" is smaller and flat. With threshold value increasing, average number need retrieving delta files increase faster with "Multi-base-states and single class delta files" than "Multi-base-states and bi-classes delta-files" as well as "Multi-base-states and multi-classes delta files". The difference of delta files number between "Multi-base-states and single class delta files" and "Multi-base-states and multi-classes delta-files" or "Multi-base-states and bi-classes delta-files" become greater. (4) "Multi-base-states and multi-classes delta-files" increases class level of delta files without increasing the number of delta file comparing with "Multi-base-states and bi-classes delta-files", but its average number to retrieve of delta files decreases a lot. That indicates "Multi-base-states and multi-classes delta-files" is a better storage method of spatio-temporal data due to its high speed to recover historical data without increasing storage space.

We can see from Table3 that the method of form a in Figure2 ("Multi-base-states and multi-classes delta files") flows the shortest path, form a in Figure2 ("Multi-base-states and bi-classes delta-files") is next to it, and form e in Figure1 ("Multi-base-states and single classes delta-files") is the longest. It also proves that "Multi-base-states and multi-classes delta-files" is a better method to store and retrieve spatio-temporal data.

Table2 proves that "Multi-base-states and multi-classes delta-files" is a better method to store spatio-temporal data in path to flow every time (from t_0 or t_{16} to t_1~t_{15}) when threshold = 15.
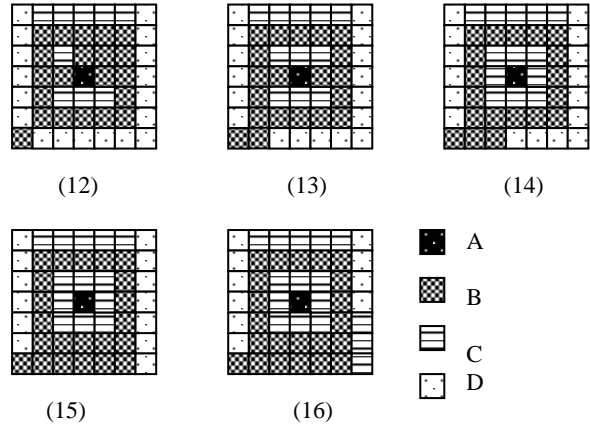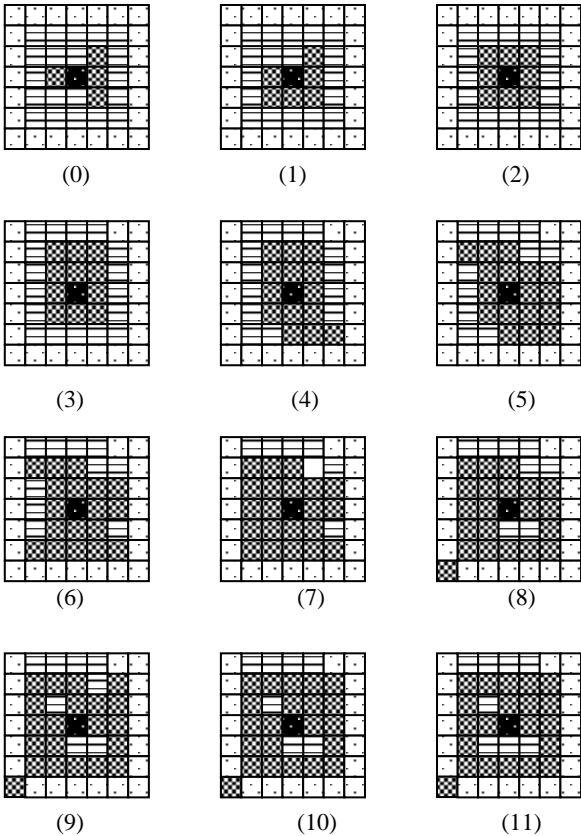
3. EXPERIMENTATION AND CONCLUSION

(0)   (1)   (2)

(3)   (4)   (5)

(6)   (7)   (8)

(9)   (10)   (11)

(12)   (13)   (14)

(15)   (16)

A
B
C
D

Figure 4. variance graph of urban land using

column row	1	2	3	4	5	6	7
1	D	D	D	D	D	D	D
2	D	C	C	C	C	C	D
3	D	C	C	C	B	C	D
4	D	C	B	A	B	C	D
5	D	C	C	C	B	C	D
6	D	C	C	C	C	C	D
7	D	D	D	D	D	D	D

Table 4. Image data at T_0

column row	1	2	3	4	5	6	7
1	D	C	C	C	C	D	D
2	D	B	B	B	B	B	D
3	D	B	C	C	C	B	D
4	D	B	C	A	C	B	D
5	D	B	C	C	C	B	C
6	D	B	B	B	B	B	C
7	B	B	B	B	B	B	C

Table 5. Image data at T_{16}

No. of delta files	Now	Column	Before changing	After changing
1	5	3	C	B
1	5	4	C	B
2	3	3	C	B
2	3	4	C	B
3	2	3	C	B
3	2	4	C	B
3	2	5	C	B
3	1	2	D	C
3	1	3	D	C
3	1	4	D	C
3	1	5	D	C
4	6	4	C	B
4	6	5	C	B
4	6	6	C	B
5	3	6	C	B
5	4	6	C	B
5	5	6	C	B
5	2	2	C	B
5	2	5	B	C
6	5	6	B	C
6	6	2	C	B
6	6	3	C	B
7	3	2	C	B
7	4	2	C	B
7	5	2	C	B
8	7	1	D	B
8	5	4	B	C
8	5	5	B	C
8	5	6	C	B
9	2	6	C	B
9	3	3	B	C
10	2	5	C	B
11	5	3	B	C
12	1	6	D	C
13	3	4	B	C
13	3	5	B	C
14	4	3	B	C
14	4	5	B	C
15	7	4	D	A
15	7	5	D	A
15	7	6	D	A
16	7	7	D	C
16	5	7	D	C
16	6	7	D	C

Table 6. Values of first class delta files

Delta file	Row	Column	Before changing	After changing
$t_4 \to t_8$	2	5	C	B
$t_4 \to t_8$	2	6	C	B
$t_4 \to t_8$	1	6	D	C
$t_4 \to t_8$	3	3	B	C
$t_4 \to t_8$	5	3	B	C
$t_{12} \to t_{16}$	3	4	B	C
$t_{12} \to t_{16}$	3	5	B	C
$t_{12} \to t_{16}$	4	3	B	C
$t_{12} \to t_{16}$	4	5	B	C
$t_{12} \to t_{16}$	5	7	D	C
$t_{12} \to t_{16}$	6	7	D	C
$t_{12} \to t_{16}$	7	2	D	B
$t_{12} \to t_{16}$	7	3	D	B
$t_{12} \to t_{16}$	7	4	D	B
$t_{12} \to t_{16}$	7	5	D	B
$t_{12} \to t_{16}$	7	6	D	B
$t_{12} \to t_{16}$	7	7	D	C
$t_8 \to t_{16}$	1	6	D	C
$t_8 \to t_{16}$	2	5	C	B
$t_8 \to t_{16}$	2	6	C	B
$t_8 \to t_{16}$	3	3	B	C
$t_8 \to t_{16}$	3	4	B	C
$t_8 \to t_{16}$	3	5	B	C
$t_8 \to t_{16}$	4	3	B	C
$t_8 \to t_{16}$	4	5	B	C
$t_8 \to t_{16}$	5	3	B	C
$t_8 \to t_{16}$	6	7	D	C
$t_8 \to t_{16}$	7	2	D	B
$t_8 \to t_{16}$	7	3	D	B
$t_8 \to t_{16}$	7	4	D	B
$t_8 \to t_{16}$	7	5	D	B
$t_8 \to t_{16}$	7	6	D	B
$T_8 \to t_{16}$	7	7	D	C

Table 7. Multi-classes delta files values

Figure4 is variance graph of land using in a city, Table4 and Table5 are two base states values. Table6 and Table7 are single class delta files values and multi-classes delta files values respectively. Based on Table4 to Table7, the flow arithmetic offered by Table3 can calculate image values at any time.

Five storage forms on spatio-temporal data based on base state with amendment are discussed. The paper proposed "Multi-base-states and multi-classes delta-files" and "Multi-base-states and bi-classes delta-files", and tested the advantages of the method to propose.

Comparing with existed methods, the two method is of save storage space, high speed to recover historical data and old-line unaffected. "Multi-base-states and multi-classes delta-files" is of higher speed to recover historic data without increasing storage space comparing with "Multi-base-states and bi-classes delta-files". "Multi-base-states and multi-classes delta-files" is realized with some raster data.

REFERENCES:

Armstrong, M. P., 1988.Temporality in spatial databases. Proceedings: GIS/LIS'88, 2, pp. 880-889.

Langran G. and Chrisman N. R., 1988. A framework for temporal geographic information. Cartographica , 25(1),pp.1-14.

Zhang Z. and Huang M., 1996. Temporal GIS Data Structure Discussion[J]. Bulletin of Surveying and Mapping, (1), pp.19-22.(in Chinese)

Cao Z. and Liu Y., 2002. An Object-oriented Spatio-temporal Data Model. ACTA GEODAETICA et CARTOGRAPHICA SINICA ,31(1),pp.87-92.(in Chinese)

Liu R. and Liu N., 2001. Extension of Spatio-temporal Data Models of Base State with Amendments and Its Implementation in Land Registration Management Systems. ACTA GEODAETICA et CARTOGRAPHICA SINICA, 30(2), pp.168-172.(in Chinese)

YU F., 2003. The Study on Spatio-temporal Data Model and Its Application in Spatio-temporal Cartographic Visualization. Beijing: China University of Geosciences(Beijing). (in Chinese)

Thomas O. and Frank S., 2001. Time-Integrative Geographic Information Systems: Management and Analysis of Spaio-Temporal Data. Springer.

Kennedy-Smith, G.M., 1986.Data Quality: A management philosophy, proceedings of Auto Carto London, pp.381-390.