THE CONSTRUCTION AND APPLICATIONS OF MAIN RIVER ENTITIES DATA

Liu Dongqin*, Su Shanwu, Lin Zongjian

Chinese Academy of Surveying and Mapping

KEY WORDS: river entities, river rank, The STRAHLER-Rules, STRAHLER-rank, location matching, Geodatabase

ABSTRACT:

A river is presented by series arcs which are spatially end in end in a river geographic database traditionally. In order to better use the data, the main content of the paper is presenting the method of how to make the separated series arcs that present the same river (river entity) into a whole entity format. The main process includes the followings. Firstly we gain the name and code information of 1:1000000 main river data by matching the river data between 1:250000 and 1:1000000 database. Secondly we calculate the STRAHLER-rank(order) of each river arc according the Strahler rules except those whose rank are higher than 5 or those who have names. Based on foregoing work we have finally carried out the task that each river has one and only identity code. The paper also shows the advantage of using the result mentioned above.

1. INTRODUCTION

A river, road or other linear geographic entity is usually presented by series arcs which are spatially end in end in our NSDI geographic database traditionally. From these data, we can not know which of those arcs compose a geographic entity, except additional specific property information.

As for river data in our NSDI geographic database, a river may be divided into the river sections (several of arcs). If the user wants to use the river data, he can only operate these geometric objects but not the river-entities.

Some data model such as National hydrological data of the United States has been set up the natural flow of water model, and can be used for water network traffic analysis, overlapped other relevant information, geo-positioning, while at the same time for the production of a variety of different types of maps and data provide the basic data update.

Therefore, we hope to improve the way of presenting the river entities in order to describe them much suitable for application, and thus it is more convenience for the users, and facilitates analysis of the data mining applications and consistency updating, etc.

The main content of the paper is presenting the method of how to make the separated series arcs that present the same river (river entity) into a whole entity format.

2. THE KEY TECHNOLOGY OF BUILDING THE RIVER ENTITIES DATA

Currently, the 1:1000000 river data of National Foundation Geographic Information Database does not have a Name or Rank coding attribute, yet. The attribute has only GEO_CODE, which defines the type information of river, such as double line river, seasonal line river etc.

The 1:250000 river data has classification attribute (in a larger scale database also has the river classification). The classifying way is taking the major rivers in the nationwide as target for encoding, referencing "Chinese mountain stream data map" of more than 1600 rivers, based on the reference watershed map and related information, General considering the drainage area, river length, flow, shipping Level, as well as the importance of the density, the rivers are divided into six grades (the sixth rank is the rivers that have the name but no classification property). The rank and name information are respectively stored in the field of CODE and NAME. The others are all unclassified.

So first we use the 1:250000 river data in NSDI to tell apart the rivers of the 1:1000000 which have CODE or NAME, thus, the up sixth level rivers are tolled apart, and then take certain method to proceeding classification calculation. Based on the classification result, we can identify each the river entity. We use the 1:250000 river data in NSDI to tell apart the river of the 1:1000000 which order is above rank 5 or which has name.

Taking into account that up sixth level rivers have been classified with the name or rank code， in order to uniform standards, we only take the no-rank rivers as the deposing and constructing entity objects.

From the existing Data, the required work is: the matching of the river, computing STRAHLER-rank of non-hierarchical river lines, and finally carries out river computing entity.

The main processes and technology are presented as following.

2.1 river matching

Because the same river has the fixed spatial location, even it is in different database, the spatial location is fixed, so according the spatial location relationship information we can match the same river in different database.

Here the purpose of spatial location matching is to identify the rivers in 1:1000000 database whose corresponding data in the 1:250000 database, which have the name and code.

We have adopted the 1:250000 rivers data of National Foundation Geographic Information Database as the basis for

* Corresponding author. dongqin@yahoo.com

the identity classification attribute of the river, and then use spatial location matching method to obtain the corresponding river attribute information in the 1:1000000 data.

Because 1:1000000 data stored as latitude and longitude form in the database, and in the process of location matching we will compare the length of the river, so it must first be converted to projection data. The mathematics base is LAMBERT projection, reference ellipsoid is Krasovsky ellipsoid, and the length unit is meter; the projection parameters adapt the uniform national projection parameters.

Through spatial location matching, we get name and code attributes of rivers in 1:1000000 database.

Fore-and-aft river matching, the river attributes can be seen in the table as follows.

Fnode_	Tnode_	Length	Geo_code
22685	22278	13550.06729	21011
22278	22034	9127.18613	21011
21446	21621	11190.24985	21011
25613	25460	3949.78915	21011
25570	25973	16022.08564	21011
25581	25431	8473.19957	21011
24449	24482	5506.39203	21014
25727	25431	7592.72353	21011
21673	21642	4408.63315	21014
21925	21707	6712.04333	21011
21599	21584	11344.67382	21011
[illegible]	[illegible]	[illegible]	[illegible]

Table 1. the attribute of the one of 1:1000000 river data

Fnode_	Tnode_	Length	Geo_code	Hydc	Name
22797	23081	7256.74640	21011	JD0006	卓摸曲
23004	23097	5006.72719	21011	JD2235	雪荣藏布
21987	21924	6114.19103	21011	JD2504	徐达曲
21666	21708	2121.30483	21011	JD2504	雄曲
21924	21759	7653.04160	21011	JD0006	乌树弄曲
25473	25378	9390.64660	21011	JD0006	司马朗曲
21936	21931	1761.97064	21011	JD2204	色荣藏布
21441	21426	3397.00584	21011	JD0006	桑曲
21989	21760	6740.29078	21011	JD0006	窃吾曲
21982	21673	8903.19462	21011	JD0006	普曲
22882	22951	2098.71882	21011	JD2404	娘曲
23280	23408	4259.24010	21011	JD2404	娘曲
24600	25226	18350.75265	21011	JD0006	念曲
25431	25282	4558.83606	21011	JD0006	泥曲
25517	25532	2567.89181	21011	JD2404	尼洋河
22699	22700	3768.76293	21011	JD0006	尼都藏布
25226	25192	2889.25015	21011	JD2245	墨竹曲

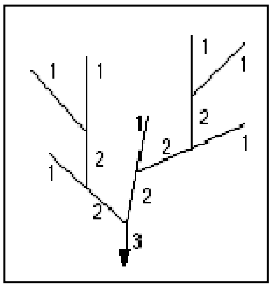
Table 2. the attribute of the 1:1000000 river data after matching with the 1:250000

The fields attribute in the tables：Fnode、Tnode、Length、Geo_code、Hydc、Name Describe Respectively the start node, the end node, length, the classify code, river code, river name

From the two tables we can see, the rivers which are upgrade 6 have got rank-code and name, and for those who have no rank we take the STRAHLER rule to compute their ranks.

2.2 Computing the STRAHLER rank of the rivers

The basic idea of the STRAHLER rule is that all primary rivers are set to level 1, while the river level for the next assignment, it must be when two or more than two intersecting arcs who are of the same grade level , the STRAHLER level of the lower reaches of the river arc are for the current level add level 1; otherwise rivers such as the intersection of the arcs are not the same , the STRAHLER level of the lower reaches will get from the highest STRAHLER level of the intersection arcs.



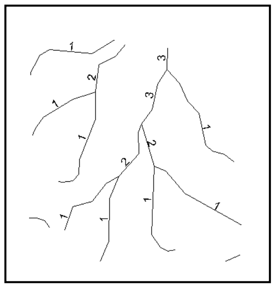


Figure 1. Figure 2

Figure 1. the chart of the STRAHLER rule

Figure 2. the computing result based on STRAHLER rule

To get right STRAHLER rank need the river has right direction flow, otherwise there will be a miscalculation. By inspection we found unreasonable classification because of the wrong river direction flow, and we have custom a river flow flip tools to correct its flow, at the same time to correct its STRAHLER rank.

STRAHLER rule classification method is most suitable for tree shape rivers classification; the calculated results show that the western region of country the tree shape rivers classification well, but for the mesh river it is not so effective, so this part should customize interactive tools, the use of human-computer interaction approach to achieve classification.

The determinations of river entities are based on the calculation result of STRAHLER rank.

2.3 Entity of the river

Entity of the river is that the same river has the only marked value in identify property field, and for the same river this field are given the same value. In this time the identify property item field named "FeatureId". Different rivers their “FeatureId” have different values.

Before computing river entities it must be thought about the problems because of the reasonable presence of pseudo-node. Here mainly exist two situations: first, two lines connected to the river, but the different part has a different property, such as reservoirs or ditch, two above of different property arcs, we would not remove different property as cost to delete the pseudo-node and append two river line; the second, a number of tiles append, the edge of the maps has pseudo-nodes. In order to ensure entities of data have consistency before and after calculation, and can link with the resource river database, we don’t deal with pseudo-node; it’s also can reduce the workload.

In accordance with the Entities method, the arcs whose STRAHLER rank is 1 need pseudo-node treatment. Get the FeaturedID of the headstream of the river; call the down flow whose STRAHLER rank is also 1 the same FeaturedID value as that of the headstream. Add up the length of the river that has the same FeaturedID value, take this length as the total length of the river. The total length will take a reference to determine which the main headstream of the river is when there are more than two headstreams.

The first steps of the Entities calculation are starting from arcs whose STRAHLER rank is 2. According to the STRAHLER rules, the upper flow of the river whose STRAHLER rank is 2, that mean two situation: the existence of two or more rivers at the headstream or have one river whose STRAHLER rank is 2, also can with some river whose STRAHLER rank is 1. Compare the length of the source of the river, and achieved the greatest length river as the headstream of the river, call the river's FeaturedID attribute value as that of the headstream'. Add up the length of the river to the total length. Stop when the entire river whose STRAHLER rank is 2 has finished computing. And then computing from arcs whose STRAHLER rank is 3; repeat the above steps until all the rivers in the calculation of the end line entity.

Generally speaking, the higher rank upstream, the longer the river length. Thus higher rank of the river line direction is the mainstream of the river; there is also the situation that the lower rank of the river line direction is the mainstream of the river direction.

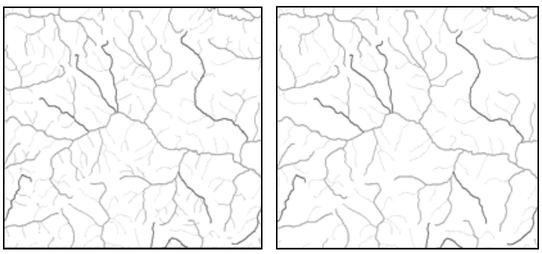
3. THE RESULTS AND APPLICATION

Through the above-mentioned work, the main river data nationwide have all entities. Finally the results are every river has its only marked. Some rivers of the results of entities as follows, according to its label can be a variety of analysis, such as geographic analysis, statistical computing, and so on.

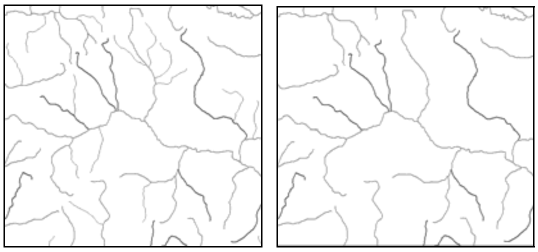
Shape	Hyd#	Hyd-id	Geo_code	FeatureId	Sum_Length	FCount
PolyLine	1947	1946	21011	83874	90594.47	10
PolyLine	2083	2082	21011	84029	46261.88	4
PolyLine	2219	2218	21011	84188	43418.63	2
PolyLine	2976	2975	21011	85057	41762.46	6
PolyLine	3009	3008	21011	85095	36493.87	5
PolyLine	2345	2344	21011	84335	35508.94	4
PolyLine	2526	2525	21011	84542	32562.65	4
PolyLine	2123	2122	21011	84076	29862.05	3
PolyLine	2062	2061	21011	84007	29741.58	4
PolyLine	1965	1964	21011	83893	29715.83	2
PolyLine	2142	2141	21011	84097	28619.16	4
PolyLine	2791	2790	21011	84831	28547.01	3
PolyLine	1730	1729	21011	83640	28241.32	1
PolyLine	2114	2113	21011	84065	27767.60	1
PolyLine	2050	2049	21011	83993	27644.42	5
PolyLine	2016	2015	21011	83954	26651.97	3
PolyLine	1891	1890	21011	83812	25853.73	1
PolyLine	3092	3091	21011	85187	25711.44	3
PolyLine	2230	2229	21011	84199	24195.71	3
PolyLine	3122	3121	21011	85222	23735.49	3
PolyLine	1670	1669	21011	83575	22274.36	1
PolyLine	2135	2134	21011	84090	21727.43	1
PolyLine	2737	2736	21011	84770	21395.75	2
PolyLine	1894	1893	21014	83815	21150.36	4
PolyLine	1839	1838	21011	83758	20577.63	1
PolyLine	4846	4844	21011	87281	20028.12	1
PolyLine	2563	2562	21011	84587	19996.03	2
PolyLine	2042	2041	21011	83984	19733.16	1
PolyLine	2331	2330	21011	84318	19443.99	3

Table 3 the example of length statistics

The 4 tiles in Figure 3 show the example of the multi-scale display of the entity river.



A B



C D

Figure 3. the multi-scale display of the entity river
A shows all the river
B shows the river whose length longer than10km
C shows the river whose length longer than20k m
D shows the river whose length longer than30km

4. CONCLUSIONS

Ways of river entities mentioned above can be applied to other types of medium and large-scale of water entities. Now in NSDI data the following six rivers have not classification coding, we can use this method of classification and entities, and then encoded according to the actual situation formulate the rules, so that water data gradually transfer into a database-oriented entities database, analytical applications and user-friendly.

Through the practical, we believe that the research of entity data can improve and effective application of NSDI data in the national economy, optimize the database, and fit for sharing and consistency updating of multiple spatial databases, but also for the future establishment of a comprehensive entity-oriented geographic database lays a good foundation.

REFERENCES

[1]Liudongqin, Sushanwu. The technology of Using ARCVIEW/ARCINFO to achieve water entities, [A]：ESRI Users volume[C]. Beijing.2004.

[2] Liudongqin, Sushanwu. The location matching of multi spatial database[J]，the science of surveying and mapping，2005.02.

[3]Shentao, Sushanwu. The research of river multi-display based on DEM [J]，the science of surveying and mapping，2003.02.

[4]Jiang jingtong, Liu Ruomei. The research of the code and name national [A], The research volume of resource and environment [C]. Beijing: The publishing house of surveying and mapping，1990.

[5]Guna, Zhouxing. Introduction of the National Hydrograph Dataset of the Unite State [J]，the science of surveying and mapping，2003.12.

[6]USGS, National Hydrograph Dataset[EB/OL]. http://nhd.usgs.gov/.

[7] Strahler A N. Quantitative Analysis of Watershed Geomorphology [J], *Transactions of the American Geophysical*

Union，1957 Vol.8, Number 6.

[8]F P da Costa, M. Grinfeld, J A D Wattis. A Hierarchical Cluster System Based on Horton-Strahler Rules for River Networks [EB/OL].
http://www.mathnet.or.kr/papers/ist/co_1.pdf.

[9]ESRI ,Using Avenue Customization and Application Development for ArcView [Z], printed in the USA,1990.

[10]ESRI, ARCINFO User's Guide [Z], printed in the USA, 1983.