

Spatial Map Data Share and Parallel Dissemination System Based on Distributed Network Services and Digital Watermark

Dong Zhang^{1, 2}, Depei Qian¹, Weiguo Wu¹, Ailong Liu², Xuewei Yang², Pen Han²

¹Department of computer science, Xi'an Jiaotong University, Xi'an, China
Navgrid@163.com

²Xi'an Research Institute of Surveying and Mapping, Xi'an, China

Abstract. A distributed map data parallel dissemination system, which can be used to distribute spatial map data to remote users through different networks, is presented. The paper first defines spatial map data, associated metadata and network resource nodes with graphic formalization definition. Then, the relations between map data and network resources are established. Based on formalized definition to the whole system, a map data dissemination framework and a series of essential distributed CORBA services, conditions input methods based on three querying conditions are presented. We also explore network map copyright validating in dissemination course, and present an improved adaptive watermarking algorithm for vector digital maps whose data is the most difficult and necessary to watermark. Different from other methods, we compare the watermarked map with the original watermarked map, not with the primitive map, before extraction. The results show that compared with Direct Copy, our dissemination system can delivery spatial map data effectively. The transfer performance of CORBA services mode is almost equal to Direct Copy. The system can approach the biggest transfer rate more quickly but depends on data amount weakly. And, the improved watermarking algorithm has better robustness.

Keywords: Map data, Metadata, Network services, Watermarking

1 Introduction

The Map Data Network Dissemination system (MDND) is dissemination of Spatial Map Data (SMD) to remote users under network environment. The SMD data have such characteristics as storage in distributed form, high precision, mass data, multi-scales, which is suitable to be applied in the network environment [1]. Meanwhile, network has become the main mean to obtain SMD today and in future. The statistics data in developed countries show that more than half information search is correlative to spatial position. However, due to limited resources and developing technical means, the map data dissemination services through network lie in a relatively lower level in China. There are several technical issues lied in SMD dissemination process. They are SMD and map metadata management and application modes, network services design

and configuration, whole data security technique, network transferring course and user ends [2]. Data security involves many research subjects. In this paper, we mainly research copyright validating algorithm being essential in dissemination. As the primary copyright validating approach, watermarking technology is difficulty and emphases, especially for vector map data focused in relative research fields.

2 Review of correlative technologies

Special dissemination systems in main developed countries have been realized and, some of them in America and England have been used in real dissemination although with simple functions and narrow application fields. Not only in developed countries, but also in developing countries, map dissemination systems are researched widely. For example, Surveying and Mapping Authority of the Republic of Slovenia has done a lot of work in the field of elaboration of the reference data sets [3].

In china, however, research on special map data dissemination system supporting all map types has just begun. Although there are vast map applications cases via internet, the map data type and quantity in the cases are limited and the area is always local.

At the same time, SMD is difficult to be processed in real time and employed in real demands from remote server nodes effectively without special dissemination system. Metadata is essential in remote SMD data access [4].

A static map may exist as a predefined image or a customized image created by a CGI Server-side utility program [5]. [6] has analyzed distributed module service technology. [7] studied data sharing and visualization based on ActiveX and Dcom. [8] gave out a JAVA network map access and display engine. [9] presented a method to implement data share by using Web Service frame. As a whole, the disadvantages through above technologies to implement SMD share and dissemination lie in the following aspects, which make a special dissemination system in urgent need:

- (1) Weak for complex computation, high precision and discrete distribution;
- (2) Weak for virtual computation, 3-D visualization in SMD dissemination;
- (3) Limited data types;
- (4) Low repetitive use efficiency for software codes and resources.

In network dissemination system, digital map copyright could not be ignored. Digital watermarking is definitely one of the most successful solutions. Vector maps, the most important type in SMD, are constructed by coordinates of points, lines and polygons whose format have the characteristic of flow data not raster data [10]. The algorithms for vector maps need to consider the robustness under revisions. Therefore, most classic watermark embedding and detection algorithms could not be applied directly to vector maps [11]. Probably due to these difficulties, we know of only a few published works on watermarking of vector digital maps. [12] subdivided a vector map into some rectangular blocks of adaptive sizes according to the density of vertices and embedded each bit by displacing the vertices in a block. [13] and [14] proposed the Fourier descriptors for watermarking of graph images. [15] divided a vector map into rectangular blocks with the uniform area.

We have constructed an experimental MDND system based on CORBA Visibroker framework and an improved adaptive watermarking algorithm for vector digital maps.

3 Dissemination Method

3.1 SMD Resources in Network Environment

Definition 1: Network Resources (NR) is defined as a relax Graph, $G=(N, A)$, where N is the network nodes set classified as SMD resources nodes N_r and metadata information nodes N_i , and A is the connections set. $|N|=n'$ is the nodes number, and the i^{th} node is represented by n_i , and the $W-n_i$ is the weight of time cost for searching n_i . $|A|=m$ is the sides number, and the i^{th} side is represented by arc_i , and the $W-arc_i$ is the weight of time cost for searching arc_i .

Definition 2: R stands for network resources NR, and $|R|$ represents the resources number. We define the relations between resources r and resources nodes (n_i) as $In(n_i, r)$, $n_i \in N_r$. $In(n_i, r)=1$ represents r is on the resource node n_i .

Definition 3: SMD Metadata is classified to six classes such as MRas, MIma, MVec, MEle, MBen and MGeo, which present the metadata of raster, image, vector, elevation, benchmark and geodesy orderly. Simply they are presented as M_{sub}^i . We define content entity and element set as follows. $SVec(k,r)$ represents the element r of entity k in MVec.

$$M_{ent} = \{SRas(k,r), SIma(k,r), SVec(k,r), SEle(k,r), SBen(k,r), SGeo(k,r)\} \quad (1)$$

Definition 4: Define Metadata base M-DB as (2). Sub-databases B_M^i can be defined as (3) corresponding to M_{sub}^i and data dictionary B_i' . The database establishment rules are denoted as C_i . Here, B_i' is relative to comparison table of data, data field, and other tables.

$$BM = \{BRas, BIma, BVec, BEle, BBen, BGeo\} \quad (2)$$

$$B_M^i = C_i (M_{sub}^i, B_i') \quad (3)$$

Definition 5: The relationships among metadata-set, metadata-base and information nodes n_i : Logically, Set (B_M^i), including one or several B_M^i , is defined as a network information node $n_i \in N_r$, here $In(n_i, r)=1$, and $r = \text{Set}(B_M^i)$.

3.2 MDND and CORBA distributed services

According to method for entity and relation, define $D = (E, R')$, as follows:

Definition 6: The entity of MDND is a five-elements set. Here, U, BM, B, c, s stand respectively for users, metadata and SMD base, query condition, and network service.

$$E = \{U, BM, B, c, s\} \quad (4)$$

Definition 7: The relations in MDND are $R' = \{R_{sub}^i, 1 \leq i \leq 6\}$, R_{sub}^i represents the all relations within entities. How to design the correlative services depends on MDND entity relations. Several entity relations and services are designed as Table 1. It should be pointed out that the given services S3-S10 (S4 exception) are formalization services, which must be changed to different service contents due to different data types [16].

Table 1. Corresponding table between entity relations and distributed services

R'	Entity/ relation	Meaning	IDL definition for CORBA Service	Service content
R_{sub}^1	(U, c)	Acquisition term input	S1:UserManagment() S2:IORService()	User management service IOR service
R_{sub}^2	(c, BM)	For metadata	S3:Metallocate()	Search for metadata-bases
R_{sub}^3	(BM, s)	Metadata bases services	S4:RightPro() S5:MetaDataQuery() S6:MetaDataGet()	Data authorization service Metadata searching service Metadata acquisition service
R_{sub}^4	(BM, B)	SMD bases location	S7:Maplocate()	SMD route pointing service
R_{sub}^5	(B, s)	SMD bases distributed services	S4: RightPro() S8:DataQuery() S9:DataPreDisplay() S10:DataGet()	SMD contents searching Pre-display and confirming SMD acquisition service
R_{sub}^6	(U, B)	DownloadSMD	S11:DailyRecord()	Log service

There are also many other absolute computations in the system called absolute services and loosely connected with network, shown in Table 2.

Table 2. Absolute services table

Absolute Service in Servers (before transferred from servers)	Absolute Service in Client (after received on clients)
S12:DataCompression()	S16:DataDe-compression()
S13:DataEncrypting()	S17:DataDe-encrypting()
S14:WatermarkEmbedding()	S18:WatermarkDetection()
S15:DataEncapsulation()	S19:DataBackuping()
	S20:CoherenceChecking()
	S21:CacheManagement()
	S22:DisplayEngine()

3.3 MDND architecture based on distributed services

The dissemination system architecture based on CORBA ORB is shown as Fig. 1. Here, only a part of main data operation services, which can reflect dissemination flow, are given. Metadata and SMD are distributed in server ends. For SMD, it is logically comprised of several bases according to the data types. The SMD Data operation services include data query, pre-display, acquisition of S8-S10 and authorization service S4. Moreover, a series of complex algorithms without close relations with network operations, such as data encryption, data compression, watermarking, etc, are designed as absolute services, which are essential to dissemination. Similar to SMD, metadata is supported mainly by the data query service S5, the acquisition service S6 and the authorization S4.

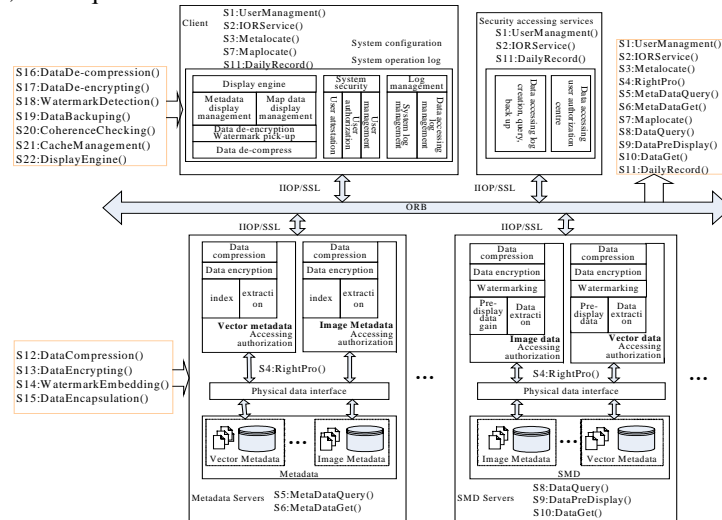


Fig. 1. Distributed CORBA services chart in MDND

The client end shows the whole dissemination flow, including logical demands presentation, data de-compression and de-encryption, watermark detection, pre-display and display of metadata and SMD, user and log management (S11, S16-22),

etc. This paper mainly research network operation services and absolute services S14, S18.

3.4 Realization of MDND

From users' demands to object data, the operation steps are as follows:

(1) Demands logical presentation: There are three query methods related to R_{sub}^1 to build correlative query condition α , they are fulfilled based on catalog tree, visualization map window, and blurry restriction condition query.

(2) Metadata and SMD retrieving services: Compared with reference [16], this paper designed metadata centralization query service Metalocate() and SMD route pointing service Maplocate(). The searching procedure can be considered as a path in graph G from start resource node n_0 to the node n_i . The dissemination flow shows as:

Dissemination (data type: Mvec,Bvec, Input: α , Mvec,Bvec, Output: SMD-Vector Map)

```
{Input Conditions( $\alpha$ );
Servers: S3:Metalocate( $N_i, \alpha$ );
If (S4:RightPro())then { S5:MetaDataQuery();
S12:DataCompression(); S13:DataEncrypting();
S6:MetaDataGet(); }
S7:Maplocate( $N_r, M_{sub}^i(k, r)$ );
If (S4:RightPro())then{ S8:DataQuery();
S9: DataPreDisplay(); S12:DataCompression();
S13:DataEncrypting(); S14:WatermarkEmbedding();
S15:DataEncapsulation();S10: DataGet(); }

Clients: {S16:DataDe-compression();
S17:DataDe-encrypting();S18:WatermarkDetection();
S22:DisplayEngine();S20:CoherenceChecking();
S19:DataBackuping();S21:CacheManagement(); } }
```

In the experiment, servers were installed in Xian Jiaotong University linked with internet, and the clients were set in Xian Research Institute. MDND is a user interactive course. In order to ensure veracity, we measured the time indexes three times by different presentation methods, and then calculated the average value.

We carried out a comparison study on transmitting performance between MDND and direct copy. Here, we used four vector map sheets of 1:50000 from Oracle database. The results are shown in Table 3. MDND can rapidly respond to network resources locating, especially for S3 and S7 services. A few instable performance indexes indicate that research institute intranet is not very steady. For mass image bigger than 100MB, transfer performance would attenuate because of repeated iterative algorithm.

Table 3. Transfer performance comparison

Map ID	F2378	F4765	F15632	F5634
Map data (MB)	19.0	22.0	17.9	18.9
S3:Metalocate() (ms)	32	15	16	16
S7:Maplocate() (ms)	16	31	16	32
S10:DataGet() (KB/s)	573	450	498	525
Copy (KB/s)	626	513	564	567

4 Copyright validating

4.1 Embedding and extraction

Embedding algorithm should explicitly make use of perceptually significant features in the data [17]. In the algorithm introduced here, the embedded intensity of watermark will be modulated adaptively according to the corresponding density of vertices. After the subdivision, according to the corresponding density of vertices, the embedded intensity of watermark in each block will be modulated adaptively as formula (5) and formula (6);

$$\alpha_s = (1 - \gamma) \times \alpha_{\max} + \gamma \times \alpha_{\min} \quad (5)$$

$$\gamma = [\max(NP_s) - NP_s] / [\max(NP_s) - \min(NP_s)] \quad (6)$$

α_s ($s = 1, 2, \dots, k * l$) is the embedded intensity of rectangle blocks, α_{\max} is the maximum of embedded intensity, and α_{\min} is the minimum of embedded intensity. α_{\max} is chosen within the range of the maximal mapping error so that the displacement won't affect visual qualities of maps. NP_s is the vertex numbers of rectangle blocks, $\max(NP_s)$ is the maximum of NP_s , and $\min(NP_s)$ is the minimum of NP_s .

According to the corresponding embedded intensity, the watermark is embedded into the vector map by displacing the vertices in the rectangles. Repeatedly, embedding each bit c times increases resiliency of the watermark against additive random noise. Formula (7) gives the method to compute c . NP is the number of vertices in the map.

$$c = \lceil NP/n \rceil \quad (7)$$

When extraction, different from other methods we compare the watermarked map with the original watermarked map, not with the primitive map. The original watermarked map, published to the users, is the unrevised watermarked map directly derived from the primitive map and has the same record order with the primitive map. If we align the watermarked map with the primitive map directly, it is very possible to make mistakes because there are two types of noises (watermark data and random noises) between the two types of map data. This has been done by minimizing the Euclid distance between vertices in the watermarked and the original watermarked map.

When extracting the watermark image, we use mean-value detection, different from the double-threshold combinative detection used by Reference [15]. The algorithm can get the valid embedding times and the total embedding bits of each pixel in the two-value watermark image through the alignment and comparison mentioned. Then the mean value of the total embedding bits of each pixel can be easily computed and converted to the corresponding pixel value of the two-value watermarking image.

4.2 Experiments and results

In the experiment, we used a 1:50000-scale vector map which has 10 feature layers and 380368 vertices, partly shown in Fig. 2, and its size is 9.85 MB. The two-value watermark image used here is a 128x128 Bmp file, shown in Fig. 2, and the size is 10.0 KB. The maximum of embedded intensity is set to 2.5m, and the minimum of embedded intensity is set to 1m. The map is divided into 30x3020 blocks. The watermark image was embedded into abscissa of vertices.

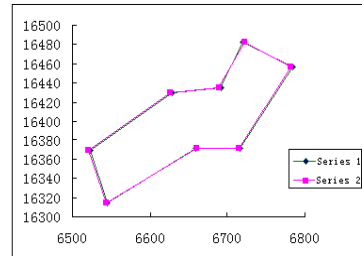
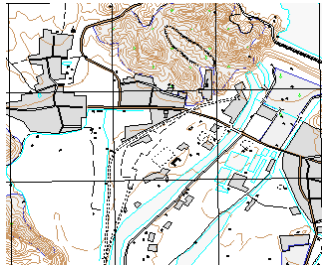


Fig. 2. Map and watermark image in experiment **Fig. 3.** Visual difference after zooming in

We compared the coordinate of watermarked map with the primitive map to observe the change of coordinate made by watermarking. The range of coordinate change is within $\pm 2m$, and the watermarked result is satisfied with mapping demand. In Fig. 3, the visual difference between the two data is still very little after zooming in many levels. Here, Series 1 denotes primitive data and Series 2 denotes watermarked data.

Table 4 shows the results of the attack experiments under random noise, vertices insertion, vertices deletion, etc.. The advantage of the improved algorithm is especially presented in random noise of amplitude 28m. Fig. 4 gives the results of

both algorithms against random noise attacks. (1) is the improved algorithm, (2) is [15] algorithm. The improved algorithm slightly outperformed the one in Reference [15] when we deleted 371527 vertices from the map data. From the experiments, we also found both algorithms withstand vertices insertion and random noise of amplitude 8m.

Table 4. Performance in various attacks

Attacks		Improved algorithm	[15] algorithm
Random noise	8 m	1.00	1.00
	28 m	0.90	0.29
vertices insertion		1.00	1.00
vertices deletion		0.98	0.96

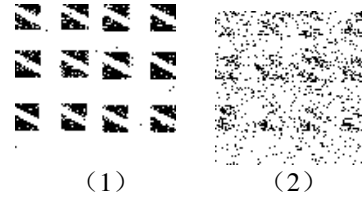


Fig. 4. Results of random noise attack

5 Conclusion and future work

The experiment results indicate that: ①Remote user can get SMD conveniently with integrated SMD and network environment, which expands data application channel. ②The services have excellent performances in metadata and SMD management, SMD location and accessing, and SMD transmitting, etc. ③User is able to access SMD data in real time only after registration, authorization, request, download, and configuration. ④The architecture is better for expanding servers, services, SMD resources, and overcomes other systems' disadvantage of bad expansibility. Furthermore, the system supports C++ environment, complex SMD computation and mass data. At the same time, an improved watermarking algorithm for vector digital map presents the concept of original watermarked map in extraction. Experiments show that the improved algorithm is more resilient against many types of attacks.

In the future, we will study user dynamic customization technology and the re-developing capability on the existing system.

References

1. Tsou, M., Battenfield, B.P, "Client/Server components and Metadata objects for distributed Geographic Information Services", Proceedings of GIS/LIS'98, November 1998, pp. 590-599.
2. Joe Futrelle, "Developing metadata standards for science data reuse in NCSA's distributed Grid architecture", 0-7803-6359-0/00, IEEE, 2000, pp. 1217-1219.
3. Irena Azman, Tomaz Peter, "Data dissemination and pricing policy for spatial data at the Surveying and Mapping Authority of the republic of Slovenia", 8th EC-GI & GIS work shop ESDI-A work in Progress Dublin, Ireland July ,2002, pp. 3-5.
4. Robert G.Cromley, Patrick McGlamery, "Integrating spatial metadata and data dissemination over the internet", IASSIST Quarterly Spring, 2002, pp. 13-16.

5. Kobben, B, "Publishing maps on the web", Chapter 6 in Web Cartography, M-J.Kraak and A.Brown(eds.), New York: Taylor and Francis, 2001.
6. Yufei Wang, Linlin Ge, Chris Rizos, et al. "Spatial data sharing on GRID", Geomatics Research Australia, No.81, December 2004, pp. 3-18.
7. Zhang De, "Map data visualization browser technology on network environment", Zhengzhou: Institute of Surveying and Mapping, Information Engineering University, 2005.
8. G. v. Laszewski, I. Foster, J. Gawor, and P. Lane, "A Java Commodity Grid Kit", Concurrency and Computation: Practice and Experience, vol. 13, Issue 8-9, 2001, pp. 643-662.
9. Wang, F, "A distributed Geographic Information System on the common object request broker Architecture (CORBA)", Geoinformatic, 2000, pp. 89-115.
10. Obelheiro P R, Fraga J S, "Role-based Access Control for CORBA Distributed Object System", Proceedings of the 7th International Workshop on Object-oriented Real-time Dependable Systems, 2002.
11. Ho-Hsun Chung, Tsuhan Chen, Kou-Sou Kun, "Watermarking 2D/3D Graphics for Copyright Protection", Proceedings of IEEE International Conference Acoustics, Speech, and Signal Processing, Vol.4, 2003, pp. 720-723.
12. Ryutarou Ohbuchi, Hiroo Ueda, Shuh Endoh, "Robust Watermarking of Vector Digital Maps", Proc ICME2002, Lausanne, IEEE Computer Society Press, Vol.1, 2002, pp. 577 - 580.
13. Nikolaidis, N. Pitas, I. Solachidis, V, "Fourier descriptors watermarking of vector graphics images", Proceedings of XIII Brazilian Symposium on Computer Graphics and Image Processing, vol. 3, 2000, pp. 9-12.
14. V. Solachidis, N. Nikolaidis, "Watermarking Polygonal Lines Using Fourier Descriptors", IEEE 2000 ICASSP, 2000, pp. 1955-1958.
15. LI Yuan-yuan, XU Lu-ping. "Copyright protection of the vector map using the digital watermark". Journal of Xidian University, Vol.31(5), 2004, pp. 719-723.
16. Klaus Krauter, Rajkumar Buyya, Muthucmaru Maheswarn. "A taxonomy and Survey of Grid resource management systems for distributed computing", Software Practice and Experience, 32(2), 2002, pp. 135-164.
17. M. Kutter, S.K. Bhattacharjee, T.Ebrahimi. Towards "Second Generation Watermarking Schemes". IEEE 1999 ICIP, Vol.1, 1999, pp. 320-323