RESEARCH ON THE SECURITY OF NATIONAL GEOSPATIAL DATA CLEARINGHOUSE BASED ON ASP

Zhang Li
Gong Jianya
Zhu Qing

KEY WORDS active server pages (ASP); national geospatial data clearinghouse (NGDC); geographic information system (GIS); Internet

ABSTRACT On the basis of the authors' experiences of setting up an NGDC Web site, this paper attempts to present some significant aspects about the security of NGDC based on ASP. They include data storing, database maintenance, new technical support and so on. Firstly, this paper discusses how to provide the security of data which is saved in the hosts of NGDC. The security model of “Networks-DB Server-DB-DB Object” is also presented. In Windows NT Server, Internet Information Server (i.e., IIS) is in charge of transferring message and the management of Web sites. ASP is also based on IIS. The advantages of virtual directory technique provided by IIS are emphasized.

An NGDC Web site, at the Research Center of GIS in Wuhan Technical University of Surveying and Mapping is also mentioned in this paper. Because it is only an analogy used for case study, the transmission of digital spatial products is not included in the functions in this NGDC Web site. However, the management of spatial metadata is more important and some functions of metadata query are implemented in it. It is illustrated clearly in the functional diagram of the NGDC Web site.

1 Introduction

Needless to say, it is very important for most GIS users to acquire and integrate the geospatial information from various districts. However, the current situation of geospatial information production and dissemination in the world is still unsatisfactory. On one hand, users do not know where the geospatial data files are stored and what geospatial data is useful for their applications, or have not necessary computer facilities. On the other hand, due to the lack of coordination and cooperation, the duplication of geospatial data production widely exist. Most of geospatial information is stored by different organizations including governmental organizations, commercial companies. What's more, the lack of geospatial data exchange and sharing mechanism results in relative low benefit of geospatial data use. It is difficult for some producers to get necessary information from other producers to integrate with or to update their own databases. In short, the value of geospatial information has not been shown exactly in GIS industry of China.

It is obvious that the information distribution technique based on Internet can play a great role in GIS industry. National Geospatial Data Clearinghouse (NGDC) is a distributed network of geospatial data producers, managers, and users linked electronically (Clinton, 1994). Through Internet, GIS
users will be able to query what geospatial data is being produced, how about its quality, where it is produced, and how to get the geospatial data economically and conveniently.

2 NGDC and ASP technology

As mentioned above, NGDC is a geospatial information distributed network system which is concerned with geospatial data producers, managers and users. So the relationship among them must be harmonized. The NGDC provides the service of geospatial information through Internet. In detail, it will allow various data formats to exist in this opened geospatial information service system and it supports the share and query of the geospatial data from different sources. The main mission of NGDC is to offer a means of fast, efficient, safe, economical service of geospatial data provision to users. At the same time, it will offer means for data providers to advertise their new products and collect users' demands and feedbacks in order to promote the geospatial data production.

To date, the model of NGDC is usually described as a provider-oriented model. In this model, every geospatial data provider is linked with Internet as an NGDC node. Users access NGDC nodes through Internet and browse the catalogues of geospatial data stored in NGDC, then they query the metadata about the available products for their applications. After selecting the desired data set, the user can send an order to the relevant producer on-line or by E-mail system. If users can not find the geospatial data available in this NGDC node for their applications, they will be able to access other NGDC nodes.

So the construction of NGDC is concerned with the planning and maintenance of dynamic Web sites linked with Internet. Since Active Server Pages (ASP) came out with its peculiar characteristics several years ago, which is applied to the construction of more and more dynamic Web sites in the diverse fields. In comparison with common gateway interface (CGI), ASP is more effective and flexible as a server scripts environment.

With HTML pages, script commands and ActiveX components, ASP can set up dynamic, interactive and efficient Web server programs. It is not important whether browsers can run those ASP codes, because all of ASP programs including scripts plugged in HTML, such as VBScript, JScript, are executed in servers. ASP programs will send a series of commands to the script engine, then the script engine translates the commands into some codes which can be executed by servers. After running the executive codes, the results will be sent by servers to users’ browsers in HTML. In this way, it is sufficient for browsers to have basic function of browse. As a result, the speed of the system increases rapidly.

NGDC Web site provides users with a catalogue of geospatial data entity, data entity and the relevant metadata. Therefore it is inevitable to access various databases in the construction of NGDC. It is convenient to connect database systems with ASP plug-in ActiveX components, so Web pages can be linked to all kinds of databases which provide ODBC interfaces for other programs. Fig. 1 shows three levels of Client/Server structure for NGDC. ActiveX components provide the objects whose tasks are to finish certain functions. So ActiveX components are of great significance in setting up

![Fig. 1 Three levels of Client/Server structure for NGDC](image-url)
3 Research on security of NGDC

This paper attempts to present some significant aspects about the security of NGDC based on ASP, such as data storing, database maintenance, new technical supporting and so on.

3.1 Security of data storing

The information stored in NGDC includes geospatial data, relevant metadata and catalogues of data products. The maintenance of all the information is a very hard task. Of course, the security of data storing is included in it. From the point of system maintenance, the security of data storing in NGDC is concerned with disk error-tolerance and back-up supporting.

With the rapid development of manufacturing technique of hard disk, the life-span of hard disk has been lengthened. Disk error-tolerance decreases usually the possibility of data-losing because of errors of hard disks. Take Windows NT Server for example, with NTFS recovery, sector protection and cheap disk array techniques (RAID1 and RAID5) the security of disk data is maintained in Windows NT Server. As one of the key techniques, disk mirror not only provides full data back-up, but also improves the efficiency of the entire system.

It is inevitable that some errors cannot be limited in spite of any error-tolerance system. In order to maintain the security of data, the significance of data should be assessed firstly and so should the loss of data-losing. There are three kinds of dump plans for database or data files: full data dump, increment data dump and combination of them. As in NGDC the need of data back-up depends on its significance.

3.2 Security of database maintenance

As for popular large-scale database systems such as Microsoft SQL Server, Sybase, Oracle, Informix, security maintenance is implemented by four levels of “Networks-DB Server-DB-DB Object” security model. (see Fig. 2.)

In Fig. 2, ① is network user’s login ID; ② DB user’s login ID; ③ DB user’s ID; ④ access rights of DB objects.

Every user has his network login ID and his password, with which network operating system identify the user. With the user ID and the password, users can login into network. Take Windows NT Server for example, Windows NT Server provides some security maintaining methods such as encoded password, minimum password length and so on.

In general, network cannot automatically permit its network users to access databases (DBs) in it. Thus, after a user login into network, the next step is to put forward DB available, user’s login ID and the password to DBMS. However, it is the exception in Microsoft SQL Server because there are two kinds of methods to identify DB users: Windows NT identifying security model (Network user ID can be regarded as DB user’s login ID. In other words, DBMS identifies user with his network login ID.) and Microsoft SQL Server combined identifying security model (DBMS identifies user with his network login ID or DB user’s login ID.)

Furthermore, the fact that a user can access DBS does not mean that he can automatically access databases in it. Only those users who have their DB user IDs stored in system tables in DB can access DB. On the other hand, even though a user has rights to access DB, he still cannot access DB objects (such as tables, views, storing procedures) in it un-
3.3 Security with ASP

In the environment of Windows NT Server, Internet Information Server (IIS) is in charge of distributing information and maintenance of Web sites. ASP is also based on IIS. When users access some ASP files in their browsers, the relevant ASP scripts will run in server and the results will be sent to users in Web pages.

Virtual directories are different from physical directories in hosts or servers. Network administrators may make good use of the mechanism of virtual directory in order to maintain the security. IIS supports virtual directory which plays a great role in the security maintenance of Web sites. Firstly, virtual directory conceals the information about actual directory structure. In normal browsers, users can get the path information of a certain Web page. If actual directory is used in organizing Web sites, the directory information of Web sites will be exposed to users linked with Internet. As a result, it is easy for the Web sites to be attacked by hikers. Secondly, it is convenient to transfer the WWW service from one server to another without updating the code in Web pages if there is the same virtual directory structure in two servers. Finally, when putting Web pages into virtual directories, administrators can assign different attributes to the directories. For example, in the construction of NGDC Web site, it is important to put normal HTML files and ASP files into different virtual directories. The attribute of directories in which normal HTML files are stored may be “Read” while the attribute of directories in which ASP files are stored may be “Execute”. On one hand, it simplifies the maintenance and management of NGDC Web sites. On the other hand, ASP source files will never be sent to user browsers. In other words, hikers cannot get the ASP source codes through their browsers. Thus it improves the security of ASP files.

ASP has an inner object named “Response”. Response object is in charge of delivering information to users, such as to send information to user browsers, to redirect to another URL, and to set cookie. “Expires”, which is an attribute of Response, is used to determine the time of a Web page stored in caches in user browser. If “Expires” of a Web page is assigned to 0, the Web page stored in caches will expire at once.

4. An NGDC model Web site in WTUSM

Some other security aspects in operational model, programming, management in the plan and construction of NGDC should be concerned. As an example, the construction of an NGDC model Web site is presented below in order to explain the security maintenance of NGDC in detail. On the basis of authors’ research on relevant problems, this NGDC model Web site was planned and deployed in early 1999.

As a model project, the purpose of construction of this Web site is to provide some useful experiences for other projects on NGDC. Therefore the process of geospatial metadata play a great role in this Web site. Fig. 3 shows the functional diagram of the NGDC model Web site. In this project, geospatial metadata tables comply with “1:50 000 fundamental geographic digital product metadata standards (draft)” in “1998 digital production techniques provisional specification”. There are 98 metadata items which describe basic information of products, map geographic parameters, information of aero-photographs, information of data quality, information of data distributor, etc. Those metadata items may be used in 4D products, such as DLGs (digital line graphs), DEMs (digital elevation models), DOMs (digital orthophoto maps) and DRGs (digital raster graphs). In fact, there are not actual geospatial data products stored in this NGDC model Web site. The main task of this Web site is to provide relevant geospatial metadata services, so the functions of data product maintenance cannot be found in Fig. 3.

In this model Web site, the following measures of security maintenance are included:
Geospatial metadata is stored into meta-database in Microsoft SQL Server. With "Networks-DB Server-DB-DB Object" security model in Microsoft SQL Server, the relations between user and access rights are set up. In order to simplify the problem, those two tasks are assigned to two DB users. One is a user who is the owner of DB objects. (Of course, he has all rights to access, update and delete DB objects). The other is a normal user who can only access DB objects such as tables.

While developing ASP programs in the integrate developing environment (IDE) of Microsoft InterDev, the functions may be fulfilled by script programs running either in clients or in servers: (1) Script programs, running in client browsers, are mainly to examine forms in HTML. VBScript or JScript involved in HTML files will be executed when certain events take place, i.e., a button in a Web page is pushed by users. (2) Script programs, running in servers, are to access databases, process forms in HTML and so on. The results, which are Web pages in HTML, will be sent to users' browsers. As a result, it improves the confidentiality of ASP programs and the efficiency of NGDC service system.

In the NGDC Web site, something has been done in order to improve the security of operation: a table named providers' information table is stored in NGDC to keep some useful information about relevant geospatial data providers, such as name, ID, passwords, contact methods and so on. Furthermore, some identifying information has been stored in the table. In other words, NGDC Web sites identify the data providers with their unique identifying information. The information may be a long irregular string whose length is less than 1 024. It is produced and maintained by NGDC. The providers' information table is stored in the server in NGDC. In this way, data producers provide geospatial products together with their identifying information through Internet.

5 Conclusions

In short, it is very convenient and efficient to distribute geospatial data in the NGDC nodes through Internet. On the other hand, with the development and construction of NGDC, there will come more and more challenges and problems about the security of NGDC. Obviously some researches and discussions in this field need to be further carried out.

References

1 Ren F H, Kenji Endo, Zhang S T, et al. Information infrastructure management system (IIMS) and networking GIS. In: Proceedings of Spatial Information Technology Towards 2000 and Beyond. Beijing, 1998.164~172
2 Zhang L, Xu Y Y, Gong J Y. Difficulties and corresponding methods in the management of geospatial information in China. In: Proceedings of Spatial Information Science, Technology and Its applications. Wuhan: Press of Wuhan Technical University of Surveying and Mapping, 1998
3 Xu Y Y. Conceptual design of The Chinese National Geospatial Data Clearinghouse Prototype: [ Master thesis]. The Netherlands: The International Institute for Aerospace Survey and Earth Sciences (ITC), 1996, 2, 91 ~92