DYNAMIC UPDATING METHOD OF GEOSPATIAL DATABASE WITH INCREMENTAL DATA

Z. K. Zhai 1, J. W. Liu 1 *, J. J. Liu 1, W. H. Zhao 1, Y. Gao 1, J. Che 1

1 National Geomatics Center of China, Beijing, China - liujiw@ngcc.cn

KEY WORDS: Geospatial Database, Database Update, Dynamic Update, Incremental Data.

ABSTRACT:

With the rapid development of computer technology and Geographic Information System (GIS), geospatial data are commonly stored and managed in the geospatial database. And geospatial data are widely used in government decision-making, environmental protection, scientific research, and national defence construction. The frequency of geospatial database updating is getting higher and higher due to economic and social development. In this paper, dynamic updating method of geospatial database with the incremental data is proposed, which means only changed features are updated into the geospatial database. Features that have not changed are not updated. This method can greatly improve the efficiency of database update. It has been applied into dynamic updating of national fundamental geographic information database of China. This database has been updated and published once a year. The current situation of data is improved continuously. It can provide more reliable data guarantee for national economic construction and social development. In the future, there will be further optimization of the dynamic updating method. We also hope more and more geospatial database can be updated by this method.

1. INTRODUCTION

Geospatial data is the reflection of natural, social, and cultural landscape on the Earth's surface. It stores the information that describes the location and attributes of things, including their shapes and representation. It is the abstraction of geographical entities existing in nature. It contains position information, attribute information and topological relation between them. Geospatial data is stored using basic spatial elements such as point, polyline and polygon. The most typical feature of geospatial data is position information (Chen, 1999). Geographic Information System (GIS) is a computer system that provides the capability to manage geospatial data. The system can relate previously unrelated information according to position (Gong, 2001). Geospatial data is processed with GIS software which can capture, store, and analyse these data. It is widely applied in government decision, scientific research. Besides, geospatial data also plays an important role in our daily life, which makes our life more convenient. Generally, geospatial data is stored in a geospatial database, which can manage data efficiently. Geospatial database is the core of geographic information system and the basics of geographic information science development (Wu, 2001).

Over the past few decades, computer hardware and software technology has developed rapidly, which leads to more available GIS software. Computer and database technology have been widely used in geographic information system construction. Database updating technology is also making continuous progress. It also provides technical support for geographic information science development. Due to development of earth observation technology, geospatial data acquisition ability has been greatly improved. Meanwhile, with the continuous development of economy and society, the frequency of geospatial databases updating is getting higher and higher. Therefore, the amount of geospatial data features is growing, which brings great challenges to database update. Nowadays, the main update method of geospatial database is to rebuild the database by version, which is time-consuming and labour-intensive. It is more suitable for the initial geospatial database construction. All features are involved in data quality inspection, quality acceptance, and database construction, which leads to large data redundancy, long update period, and lots of repetitive work. In the meantime, the version-specific data can only describe the instantaneous state of the database. There is no connection between different versions. It is impossible to analyse changes between different features and predict future trends of the specific feature (Chen, et al., 2004; Chen, et al., 2007). Therefore, the appearance, existence, and disappearance of each feature cannot be traced. However, with the continuous development of economy and society, various fields continue to put forward higher requirements for geospatial database update. It is urgent to shorten the database update period to meet these needs. At present, most of the geospatial databases have completed the initial database construction. The database needs to be updated at certain intervals. Besides, changed features in the geospatial database are the focus of our attention. There is no need to update features that have not changed. The number of updated features in different regions is inconsistent. We need to find features that have really changed, which traditional method can no longer meet the need. The dynamic updating method of geospatial database can greatly improve the database update efficiency. The geospatial database is updated with the incremental data, which means that only changed features are updated into the geospatial database. Features that have not changed are not updated. The amount of data submitted for database update is greatly reduced. Therefore, the efficiency of database update is increased (Wang et al., 2018; Liu et al., 2016).

* Corresponding author
This method can effectively reduce database redundancy and workload, especially for multiple units collaborate and submit updates online. The change of the geospatial database can be easily found based on incremental data. Meanwhile, spatial analysis and statistical calculation of database update can be easily performed. Besides, the life cycle of feature can be recorded. Each feature can be traced and analysed for any change. It can better meet the needs of economic and social development for geospatial data.

2. GENERAL DESIGN

The dynamic updating method of geospatial database means that the database is updated partial. The principle of this method is only updating the incremental data into old database. Features that have not changed will not be updated. In order to update the database with the incremental data, database structure needs to be upgraded. First, a unique feature identification field needs to be created for each feature in database. When incremental data is submitted, it is used to associate with the data in database. Second, update time field and update status field must be created for each feature, which is used to record the latest update time and update status information of features (Wang and Liu, 2015).

The key to this method is to update and manage each feature of the database. We need to know the time of each feature’s birth and disappear. Meanwhile, how does each feature change during its lifetime is more important. In order to manage each feature of database, data production mode and database building technology need to be redesigned (Chen, et al., 2010).

The core of dynamic updating method is to find features that have really changed during database update process, which is incremental data. There are many ways to find changed features. We can also choose features that we focus on or features in specific areas that we concern. This can greatly reduce updated data consumption. In this method, the most important information of each feature is unique value and update status. Every feature in geospatial database must be given a unique value. This value cannot be modified. This is the key to update the database (Wang and Liu, 2018).

In data production process, changed features are found and marked. During data quality inspection process, correctness of incremental data is checked. At the meantime, the incremental data must be consistent with the background data. When database needs to update, only incremental data are updated into database. Features that have not changed are not updated into database, which greatly reduces the workload of database update. Meanwhile, each feature's birth, change, and demise can be monitored by its unique value. This strategy can considerably enhance database update efficiency (Figure 1).

3. TECHNICAL METHOD

3.1 Database structure adjustment

In order to record version information and update status of each feature in the database, the database structure need to be adjusted. Based on the original data structure, three database fields must be added in order to update geospatial database with the incremental data, including database identification field, version identification field and update status identification field (Zhang, et al., 2014). Therefore, same features between different versions can be associated automatically. Latest features are stored in the latest version of the database. Changed features are stored in the historical database, which can reduce data redundancy. There is an association between historical and latest database. It is convenient for change extraction and statistical analysis (Chen and Zhou, 2008).

In addition to general fields, three fields are added to the database. The field named id is vital for dynamic updating method. Each feature in database is assigned a unique value. This value is unique in the database. The version of each feature is represented by the version field. This field records the time when the feature was created. The field named status reflects update status of the feature. This feature records the update change type of the feature (Chen, et al., 2008). The specific field names and uses are shown in the Table 1.

| field type | field name | field function |
|------------|------------|---------------|
| general fields | objectid | This field is used to uniquely identify each feature. |
| | shape | This field stores the geometry of each feature. |
| | … | … |
| database fields | id | This field stores unique value of each feature. |
| | version | This field represents the version of each feature. |
| | status | This field stores update status of each feature. |

Table 1. Functions of fields in data.

Data structure is the basis of database update. It is also the premise of database consistency. It determines the use and analysis of data later.

3.2 Incremental Data Production based on Background Database

Based on background database, changed features are found and marked according to remote sensing data and collected industry information during data production process. Incremental data are extracted. Incremental data must be consistent with the background database.

During process of incremental data production, there are many ways to find changed features. Firstly, change features can be found according to professional data that collected from other units. These data are more authoritative. Secondly, the resolution of remote sensing data is getting higher and higher. Changed features can easily found by automatic or manual means based on high resolution remote sensing images. Those changed features that are uncertain can be identified by field investigation furtherly. Thirdly, network geographic information has characteristics of good current situation and last update frequency. It can be also used for changed features discovery. The efficiency of changed features discovery can be improved greatly.
In the process of database update, three types of features need to be updated into the database, including new features, modified features and deleted features. Incremental data is consisted of these features. The database fields for these features need to be marked correctly. Otherwise, it will cause a database update error.

New features are features that generated in process of data production. These features need to be added into database. If new features are found in process of data production, status field of these features should fill in “new”. Version field should be filled in current time. For example, if a feature appears in 2021, status and version field should be filled in “new” and “2021”, respectively (Figure 2). The value of id field can be empty. This field will be assigned a value in process of database update uniformly. New features cannot overlay or conflict with features in old database. Otherwise, there will be topology errors in database.

Figure 2. Database fields information of new features

Modified features are features that have changed during process of data production. Graphics or attributes of these feature may have changed. Original features in old database need to be replaced by these new features. If modified features are found, status field of these features should fill in “modified”. Version field should be filled in current time. For example, if a feature in old database is modified in 2021, status and version field should fill in “modified” and “2021”, respectively (Figure 3). The value of id field cannot be modified. Otherwise, features in old database cannot be replaced correctly. Modified features must maintain correct topological relationship with the original features in old database.

Figure 3. Database fields information of modified features

Deleted features are features that need to be deleted. These features cannot be physically deleted directly. Status field of these features should fill in “delete”. These features need to be stored in incremental data. For example, if a feature is vanished in 2021, its status field should fill in “delete” (Figure 4). If newly added features need to be deleted, they can be deleted physically.

Figure 4. Database fields information of deleted features

After data production is finished, incremental data are extracted based on background database. Then incremental data is ready for quality inspection.

3.3 Correctness and Consistency Check of Incremental Data

After incremental data production is finished, correctness and consistency check of incremental data is essential (Zeng and Gong, 2004). The inspection contents of incremental data mainly include correctness of data integrity and data structure, especially database fields. Features that have not changed will no longer be checked repeatedly. This makes sure incremental data can be updated into background database successfully (Shang, et al., 2012; Liu, 2010).

During process of quality inspection, correctness of incremental data needs to be checked, including data structure correctness check, coordinate system correctness check, attribute correctness check, and topology correctness check (Liu, et al., 2006; Du, et al., 2012; Zhai, et al., 2021). Meanwhile, consistency between changed features and consistency with the background database also need to be checked (Figure 5).

Figure 5. Correctness and consistency check of incremental data

The process of quality inspection is very important. Besides, unique value of each feature must not be changed. Then incremental data is ready for database update.

3.4 Geospatial Database Updated with Incremental Data

When incremental data are checked, geospatial database will be updated with the incremental data. Features that have changed are updated into geospatial database. Features that have not changed will not be updated. Through unique value of each feature, relationship between changed features and updated features in database is automatically identified. Automatic update of features in database can be done.

This method is to update associated features in database according to unique value, update version and update status of
changed features. New features will be added into database directly. Version and status information of new features are recorded at the same time. For features that need to be modified, shape or attribute information will be modified according to change features. Version and status information of modified features are recorded. Meanwhile, original features will be moved into historical database. Deleted features will not be deleted physically. These features will be also moved to historical database. They will not be stored in latest database. Disappeared time of deleted features is recorded (Figure 6). The update process of geospatial database is completed, which can considerably enhance database update efficiency.

Figure 6. Flow chart of incremental data updated into database

The historical database records features that replaced in the process of database update. After overlay analysis between the latest database and historical database, historical version data can be traced back easily. At the same time, birth, change and disappearance of each feature in historical database can be analysed conveniently. Data structure, spatial reference, attribute field of historical database is consistent with the latest database. During the process of database update, features that transferred to the historical database do not need to change anything. Meanwhile, in order to trace back historical version data, an additional field is added to record the time when features disappear.

4. ENGINEERING APPLICATION

Fundamental geographic information data is a type of geospatial data. Fundamental geographic information data plays a vital role in various industries of our society. It is the important part of national spatial data infrastructure. As a unified spatial positioning framework and spatial analysis substrate, national fundamental geographic information database is a vital basic and strategic information resource for national economic construction, social development, and national defence construction (Chen, et al., 2012; Wang, 2006).

In general, national fundamental geographic information database is mainly updated in two ways: version update and incremental update. Version update means that old version data in database are replaced by new version data directly. This method is suitable when the database is first built. Incremental update means that only changed features are updated into database. Those features that have not changed will not be updated. This method is more suitable when the database needs to be updated. This method can shorten the time of database update (Shang, et al., 2012; Wang, et al., 2013).

In order to improve the efficiency of database update, data structure of original geospatial database was adjusted. Due to national fundamental geographic information database has been built, dynamic updating method of geospatial database is used to update national fundamental geographic information database (Figure 7).

National fundamental geographic information database of China has been first built in 2006. Then all features of this database are updated in the next five years. National Administration of Surveying, Mapping and Geoinformation of China launches a project to update national fundamental geographic information database through the dynamic updating method since 2012, which aims to update this database continuously and quickly (Liu, et al., 2014; Liu, 2015; Zhang, et al., 2015). Flow chart of national fundamental geographic information database update is shown in Figure 8.

Figure 7. General technical design of dynamic updating of national fundamental geographic information database update

According to dynamic updating method of geospatial database, data structure of geographic information database was adjusted. Database fields are added to data in order to update the database with the incremental data, including id field, status field, and version field. The whole geographic information database update task is divided into several parts. Each unit performs data production within their task region at the same time. In order to make the whole work more efficient, inspection check software for incremental data is developed to ensure the incremental data’s quality. We also develop processing software to update the incremental data into background geographic information database. These tools can process data automatically. There is no human intervention required. After these sub-databases finish updating, incremental data will be committed. The submitted incremental data must be consistent with surrounding data. Incremental data must be collected and extracted in strict accordance with technical regulations. All incremental data must pass the quality inspection. The import of incremental data updated into the
database is also completed through automatic software tools. Benefiting from this approach, this database is updated and published annually.

So far, national fundamental geographic information database of China has been updated ten times and will continue to be updated in the future. Compared with traditional methods, the number of updated features is greatly reduced. The workload is greatly reduced. These incremental data can also be used to update database at other scales. These data can be released in time for the people and other departments to use.

5. CONCLUSIONS AND DISCUSSION

In recent years, computer technology and remote sensing technology has made great progress. The development of geographic information system technology and spatial database has also greatly improved the performance of geospatial database. With the continuous development of economy and society, many industries also have growing needs for geospatial data. Cities are developing faster and faster. Besides, Geospatial data also plays an important role in our daily life, including navigation, travel, and shopping. The frequency of geospatial database update is getting higher and higher. There is an urgent need to develop new technology to update the database efficiently.

The method of updating geospatial database with all features is relatively mature. But this method is time-consuming. There is a lot of duplication of work. In this method, all features in old geospatial database are replaced by new features directly. It is more suitable when the database is first built. The database needs to be updated by more efficient method. Features that have not changed in database do not need to be updated. We just need to update changed features into database. In economically developed areas, many features need to be updated. There are few changed features in desolate places. Most features in these places will not change for a long time. There is no need to update these features. Therefore, we only need to update changed features in geospatial database. It can greatly reduce the workload.

Updating geospatial database with the incremental data can greatly improve the efficiency. The amount of updated data for database is significantly reduced. Important features or features in important regions that we concern can be updated individually, which is more flexible and convenient. More importantly, changed features in background database is clear and intuitive. Frequency of updated features in each region can be easily counted and mapped. Each feature can be tracked by specific database fields.

In this method, only incremental data needs to be checked instead of full data before updated into background geospatial database. Then incremental data are submitted to update background database, which can effectively reduce workload. Moreover, it can prevent mishandling of features which have not changed. Besides, it is more suitable for online collaboration. Incremental data play a critical role in database update. Based on background database, incremental data should be found and marked correctly. Consistency between incremental data and background database need to be checked carefully. Database fields in incremental data should be marked correctly. Otherwise, there will be errors in geospatial database.

However, this method still has room for improvement. Currently, most of changed features are found through manual method. The efficiency of this method is not very high, and it is easy to lead to mistakes. In recent years, big data, artificial intelligence, and machine learning technologies have developed rapidly. It has brought revolutionary changes to many industries and our daily life (Liu, et al., 2019; Ning and Wang, 2010).

Image classification technology has made a great breakthrough. At the same time, the spatial resolution of remote sensing and UAV images is getting higher and higher. How to make better use of these technologies and these data for the discovery of change features is the next work to be done. Automatic extraction and intelligent processing are the focus of future research (Figure 9). Meanwhile, this is also an important development direction of geographic information science (Xiao, et al., 2015; Li, 2016).

With the development of internet and the advantage of worldwide information revolution. Everyone can participate in the production of geospatial data. For example, maps have become a part of our daily life. We can enrich map data by marking places near us on the map. There are also huge amounts of public data on the Internet. These data are updated quickly and some are even more accurate. How to make better use of these resources to find changed features in geospatial database is also very important, which can greatly reduce the workload of our field work (Liu and Gao, 2017; Wang, 2016).

The quality inspection of incremental data can basically be processed automatically. However, some inspection algorithms need to be further optimized. Inspection tasks and processes can be further refined. How to further improve the efficiency of quality inspection software is also our further work. Moreover, cloud computing technology can also be used in our quality inspection software. Besides, with the development of mobile internet technology, how to use internet technology for cooperation is also an issue that we should consider in the future. After the geospatial database is built, how to use these data is also very important (Liu, et al., 2014 and Chen, et al., 2009).

The purpose of these data is to serve the society and our life better. In the past, most geospatial data are provided by the way of offline. Geospatial data need to be copied with hard disk and transported manually or by express. The hard disk may be damaged during transportation. Returning to fetch these data may waste a lot of time. With the rapid development of internet technology, the traditional offline service needs to be changed to a real-time online service. Geospatial data can be transported on the Internet in the future. Specific data needs to be encrypted during transmission. The application based on geospatial data is springing up like mushrooms after a rain, which is profoundly affecting the way of our work and life. It is a trend for geospatial data to provide services through the internet in the future. The database can be updated online in real time, so users can use it without any feeling. Everyone can easily use this data with their computers or mobile phones.

In general, this method has great advantages in updating geospatial database. It can greatly improve the efficiency. This also allows us to manage features in multi temporal databases.
In the future, there will be further optimization of the dynamic updating method. We will continue to optimize this method. We also hope more and more geospatial database can be updated by this method. We will continue to promote the application of data sharing services to serve the government and the people.

ACKNOWLEDGEMENTS

I would like to thank Academician Chen, DR Wang, and my colleagues for their previous efforts. The authors would also like to thank the National Geomatics Center of China for the support.

REFERENCES

Chen, S.P., 1999: *Introduction to Geographic Information Systems*. Science Press, Beijing.

Gong, J.Y., 2001: *Fundamentals of Geographic Information Systems*. Science Press, Beijing.

Wu, L., 2001: *Geographic Information System: principle, method, and Application*. Science Press, Beijing.

Chen, J., Li, Z.L., Jiang, J., Zhao, R.L., 2004: Key Issues of Continuous Updating of Geo-spatial Databases. *Geomatics World*, 2(5), 1-5.

Chen, J., Zhao, R.L., Wang, D.H., 2007: Dynamic Updating System for National Fundamental GISs: Concepts and Research Agenda. *Geomatics World*, 5(5), 4-9.

Wang, D.H., Liu, J.J., Zhang, Y.J., Liu, J. W., 2018: Thoughts on Upgrading of National Fundamental Geographic Information Database. *Geomatics World*, 25(2), 1-6.

Liu, J.J., Wu, C.C., Yang, M., Du, X., Liu, J.W., 2016: Thoughts of Demand-Response Real-Time Updating of Fundamental Geographic Information. *Geomatics World*, 23(2), 79-82.

Wang, D.H., Liu, J.J., 2015: Key Techniques for Dynamic Updating of National Fundamental Geographic Information Database. *Acta Geodaetica et Cartographica Sinica*, 44(7), 822-825.

Chen, J., Wang, D.H., Shang, Y.L., 2010: Master Design and Technical Development for National 1:50000 Topographic Database Updating Engineering in China. *Acta Geodaetica et Cartographica Sinica*, 039(001), 7-10.

Wang, D.H., Liu, J.J., 2018: Dynamic Updating Engineering Technology of National Fundamental Geographic Information Database. Surveying and Mapping Press, Beijing.

Zhang, Y.J., Liu, J.J., Liu, J.W., Zhao, R. L., Wu, C.C., 2014: Building Feature based Multi Spatio-temporal Topographic Database Management System. *Geomatics World*, 21(1), 29-32.

Chen, J., Zhou, X.G., 2008: Incremental Updating of Spatial Database Based on Topological Linkage: Taking Cadastral Database’s Updating as an Example. *Acta Geodaetica et Cartographica Sinica*, 37(3), 322-329.

Chen, J., Liu, W.Z., Zhang, J.Q., Pan, L., 2008: Research Progress of the Model and Method for GIS Database Updating. *Geomatics World*, 2008(3), 12-16.

Zeng, Y.W., Gong, J.Y., 2004: Implementing Technique of Spatial Data Quality Control and Evaluation. *Geomatics and Information Science of Wuhan University*, 29(8), 686-690.

Shang, Y.L., Wang, D.H., Liu, J.J., Zhang, Y.J., 2012: Study on Quality Control of Fundamental Geographic Information Data. *Geomatics World*, 2012(1), 13-17.

Liu, J.J., 2010: Study of Software Design for Quality Check of Fundamental Geographic Information Data. *Bulletin of Surveying and Mapping*, 2010(11), 18-21.

Liu, J.J., Wang, D.H., Shang, Y.L., 2006: Design and Development of Quality Checking Software for National 1:50000 DLG Database. *Geomatics World*, 4(4), 23-26.

Du, X., Shang, Y.L., Liu, J.J., Zhao, R. L., Li, X.M., Wang, L.X., Zhang, Y.J., 2012: The Quality Controlling on Integrated Interpretation and Surveying Updating of 1:50000 DLGs. *Geomatics World*, 2012(3), 10-12.

Zhai, Z.K., Gao Y., Liu J. J., Zhang Y.J., Liu J. W., Zhang, X. Q., Zhao W. H., Che J., 2021: Design and Development of Geospatial Data Quality Inspection Software Based on Check Rules. *Geomatics World*, 28(1), 100-104.

Chen, J., Wang, D.H., Shang, Y.L., 2012: Design and Implementation of National 1:50000 Fundamental Geography Information Database Updating Project. *Geomatics World*, 10(1), 8-12.

Wang, D.H., 2006: Research and Implementation of Construction of 1:50000 Scale Topographic Database in China. *Geomatics World*, 4(004), 4-5.

Shang, Y.L., Zhao, R.L., Liu, J.J., 2012: National 1:50000 Fundamental Geographic Information Database Updating Project. *Geomatics World*, 10(1), 6-7.

Wang, D.H., Liu, J.J., Shang, Y.L., Zhao, R.L., Zhang, Y.J., Li, X. M. 2013: Dynamic Updating of National Fundamental Geography Information Database. *Bulletin of Surveying and Mapping*, 2013(07), 1-4.

Liu, J.J., Zhao, R.L., Zhang, Y.J., Li, X.M., Liu, J.W., 2014: Updating Key Features of National 1:50k Topographic Database. *Geomatics World*, 21(1), 37-40.

Liu, J.J., 2015: Construction and Updating of National Fundamental Geographic Information Database. *Bulletin of Surveying and Mapping*, 2015(10), 1-3.

Zhang, Y.J., Liu, J.J., Liu, J.W., 2015: Technology Method of 1:50000 Topographic Database Incremental Updating. *Geomatics World*, 22(6), 126-129.

Liu, J.J., Chen, J., Zhang, J., Zhang, Y.J., Liu, J. W., 2019: Geographic Information Dynamic Monitoring in Intelligent Era. *Geomatics and Information Science of Wuhan University*, 44(1), 92-96.

Ning, J.S., Wang, Z.T., 2010: The Newest Progress of Surveying and Mapping Oriented Information State. *Science of Surveying and Mapping*, 35(5), 5-10.
Xiao, J.H., Peng, Q.S., Li, H.T., 2015: “Geomatics 4.0”: Surveying, Mapping and Geoinformation in the Internet Era. *Bulletin of Surveying and Mapping*, 2015(7), 1-4.

Li, D.R., 2016: Towards Geo-spatial Information Science in Big Data Era. *Acta Geodaetica et Cartographica Sinica*, 45(4), 379-384.

Liu, J.N., Gao, K.F., 2017: Challenges and Opportunities for Mapping and Surveying and Location Based Service in the Age of Intelligence. *Geomatics and Information Science of Wuhan University*, 42(11), 506-517.

Wang, J.Y., 2016: Create a New Era of “Mapping and Geographic Information Science and Technology Based on Internet Plus”, *Bulletin of Surveying and Mapping*, 33(1), 11-14.

Liu, J.J., Du, X., Zhao, R.L., Liu, J.W., Zhao, W.H., 2014: Design and Research on a New System of Online Field Surveying for Geographic Information Rapid Update. *Geomatics World*, 21(2), 112-115.

Chen, J., Ding, M.Z., Jiang, J., Zhou, X., Zhai, Y., Zhu, W., 2009: From Off-Line Data Provision to Online-information Service. *Geomatics World*, 7(2), 6-9.