CONCEPTUAL FRAMEWORK FOR GEOSPATIAL DATA SECURITY

Sangita Zope- Chaudhari[1] and P. Venkatachalam[2]

[1]Research Scholar, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India
[2]Professor, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India

ABSTRACT

Due to rapid growth of distributed network and Internet, it becomes easy for data providers and users to access, manage, and share voluminous geospatial data in digital form. With Internet, it becomes very easy to distribute and copy geospatial data. Therefore, it becomes necessary to protect geospatial data from illegal and unauthorized usage. The increase availability of tools and techniques to access and disseminate geospatial data has made more urgent to protect it. Security mechanism should be enforced at geospatial data storage, dissemination, and at data retrieval to protect geospatial data from illicit users. In this paper, conceptual framework for geospatial data security is proposed by considering unique security requirements of geospatial data.

KEYWORDS

Geospatial data, Watermarking, Security, Outsourcing

1. INTRODUCTION

In recent years, the advancement in computer networks and World Wide Web (WWW) have discovered many new business, scientific and social openings in the form of electronic publishing, real-time data availability and sharing, collaboration among computers, digital repositories and many more. Geospatial data can be stored efficiently and with a very high quality, and it can be manipulated very easily using computers. Furthermore, it can be transmitted in a fast and inexpensive way through data communication networks without losing quality. With digital data distribution over WWW, authentications are more endangered than unlimited copying.

Geospatial data is gathered primarily from topographic/thematic maps, global positioning system (GPS)/ ground based observations, aerial and satellite sensors. Such a digital spatial data generation involves complex and expensive efforts. A typical geospatial data repository contains several thematic layers. These layers are managed by government or local organizations depending on theme. For example, water resource thematic layer is managed by department of water resources whereas census layer can be managed by census department at state or country level. As single repository is managed and used by private and public sectors, security mechanism should be used to specify and enforce different access methods so that if user is accessing data from one layer, he should not be able to get sensitive data from other layers residing in the spatial database. Also, dissemination of geospatial data over the network is a complex and massive task in scale and dimension. Many national and international agencies are primarily focusing on coordinated development, use, sharing, and dissemination of geospatial data among wide range of

DOI : 10.5121/ijdms.2013.5504

government agencies and offices. Use of security mechanism posed at various levels namely geospatial data storage, dissemination, and retrieval of geospatial data from trusted third party could play a significant role in geospatial data environment.

Considerable work has been done to build security infrastructure for relational data and semi-structured data. Plenty of access control and security frameworks have been proposed and deployed in database systems [1, 2]. Most of them are role based access control [3, 4] or credential based access control mechanisms [5, 6]. Unlike relational data, geospatial data can be characterized by complex data objects and complex relationships between them. Securing such type of data at storage level is a challenging task and yet not fully understood and developed. Although access control mechanism in relational database is well developed, same cannot be adopted for geospatial data. Some access control policies for geospatial data have been reported. However, it considers only one type of geospatial data (either raster or vector) [7, 8]. Moreover, it only deals with read privileges and it is not sufficient for dynamic applications. Also, the subject (user) and the objects (features) used by subject are dynamically and rapidly changing. Thus, it is required to develop an access control policy which will solve these issues.

Due to voluminous nature of geospatial data, whole data could not be stored at single server. Also various agencies are dealing with different data formats. According to their usage, geospatial data needs to be outsourced on some trusted servers by applying some modifications. Although well-known geometric or cryptographic transformations [9, 10] and watermarking techniques [11, 12, and 13] have been used for secure outsourcing and distribution of geospatial data respectively, they alone cannot provide complete security. Currently, only simple transformations are being used which can be easily attacked. Therefore, development of secure and efficient scheme is required.

In the context of the above listed problems, security is not only important at storage level but also at the distribution level. It is necessary to incorporate security at geospatial data repositories, geospatial data warehouses, and at distribution and outsourcing of the geospatial data. In this paper we first define characteristics and security requirements for protection of geospatial data. In order to fulfill these requirements, we propose a conceptual framework which provides security at storage as well as dissemination level.

2. GEOSPATIAL DATA SECURITY REQUIREMENTS

Geospatial data have some distinct characteristics: It is multidimensional. Spatial position has to be defined using x, y, z, and time depending on application; It is voluminous; same data can be represented at different levels of spatial resolutions/scales; Geospatial data analysis and retrieval is reliant on how the data is represented in database; It requires many special methods for its analysis; Updating geospatial data is very complex and expensive task. Depending on the characteristics of geospatial data, various security requirements need to be established in order to make geospatial data secure on distribution network [14]. These requirements are:

- Protection of Geospatial Data regarding privacy

  It deals with unauthorized access and misuse of geospatial data. Violation of this requirement results in illegal copy of data, tampering and forgery of data, exposure of privacy information etc.

- Ensuring Confidentiality of Nondisclosure Geospatial Data

  Some high precision data is permitted to be used by military or government agencies only. Access to such data should be prevented from the users that are not authorized to access it. Failing to it may result in leakage of defense related nondisclosure information, threat to safety of residents, illegal copy of data, tampering and forgery of data, and data error.

- Ensuring Integrity and Authenticity of Geospatial Data

  This requirement deals with completeness and correctness of geospatial data. Geospatial data should be protected from unauthorized modifications. Also, geospatial data should come from authentic sources. The possible damages are influence on services which use geospatial data, maltreating tampered digital map as an official document, and trouble or crime caused by tampered geospatial data.

- Management of the Access Privilege of Geospatial Data

  Authorization policy should be enforced to specify "who" can access "what", "where", and "how". If some errors occurred in setting the access privilege of geospatial data, the possible damages are serious threat over the security of entire database and disturbance in services that use geospatial data.

- Prevention from Violation of the Copyright of Geospatial Data

  Copyright protection should be applied to geospatial data to provide authentication and origin tracing. Violation of this requirement causes infringing author profits and prosecution of legal liability relevant to protection of copyright, masquerading of an author or a source, tampering and forgery of data, and illegal copy and distribution of data.

- Ensuring Availability of Geospatial Data

  Ensuring availability means geospatial data should be available as and when required, failing to which causes bad influence on the data service, interference in response time of disaster and emergency, and failure of interoperability between systems.

3. PROPOSED CONCEPTUAL FRAMEWORK

In the last few years, geospatial data security has become one of the main areas of research all around the world. As geospatial data is important and confidential, privacy as well as security of this data is of main concern. One needs to consider characteristics of geospatial data while designing security policy for it.

Success of any security policy depends on its resilience against various malicious attacks, their threats, and safeguards achieved for them. While dealing with security of geospatial data, nature of data should be analyzed first and then depending on it, security policy has to be designed. For example, the policy designed for vector data may not fully work for raster data and needs modification.

The proposed framework includes two layers at which security is posed. Internal layer i.e. storage layer mainly deals with fine grained and flexible access control of geospatial database. Authentic data outsourcing allows transforming and outsourcing private geospatial data to trusted servers with the help of trust and privacy management. Outer layer i.e. distribution layer mostly deals with users interacting between themselves and with GIS web services or trusted servers to retrieve geospatial data. Whenever any user wants to communicate to web services or trusted servers, mutual authentication is required which will make system safe from malicious attacks by unauthorized users. The proposed conceptual framework is depicted in figure 1.

Geospatial data resources like Geospatial data repositories, satellite images, aerial photographs, maps, thematic layers, and GPS data are processed and conceptualized using object based or field based model. Geospatial data is stored in different formats at various repositories. Therefore, a common format is specified to handle interoperability. Geospatial data is modeled using object oriented or object relational database management systems.

1.1 Storage Layer

Access control policies regulate the data access. The idea is to maintain control over who has access to data via various credentials based on function and nature of data. While dealing with geospatial databases, roles of the users as well as hierarchy of spatial data are of equal importance. The authorization model for access control will include

- Subject- user id or role
- Object – basic component and relationship
- Event – denotes if any event occurs to permit/deny access permission
- Context –Contextual information like time and subject locations considered for permission granting operation
- Permission – permit/deny
- Operation- type of operation to be executed on spatial data.
- Spatial window – specified spatial region of interest

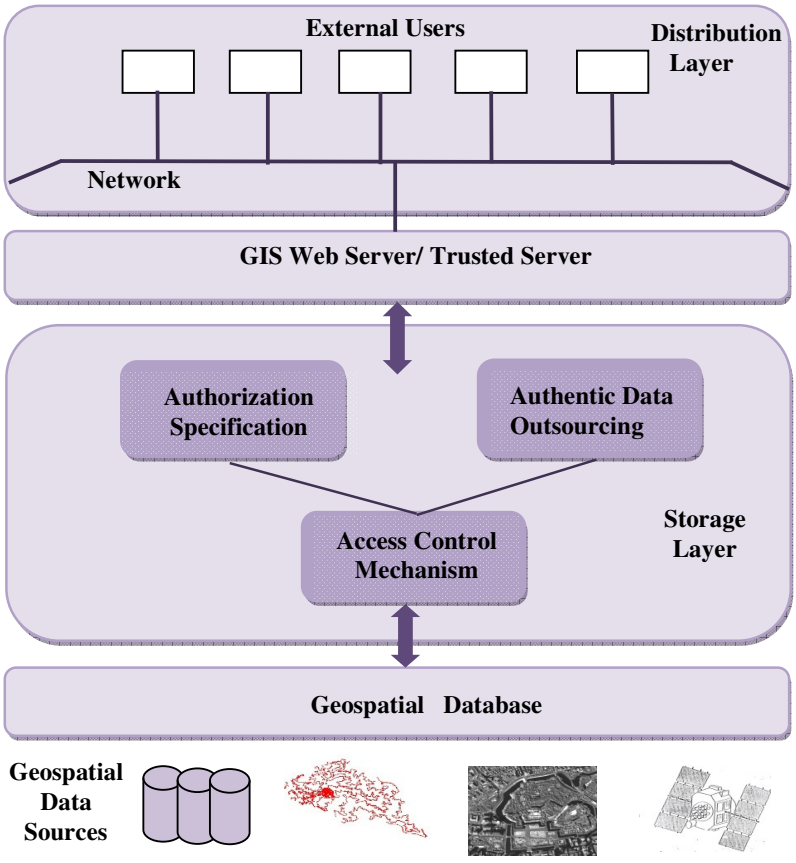

Figure 1. Conceptual framework for geospatial data security

Success of access control mechanism depends on authorization model as well as the structure of data storage at database. Mostly B+ -trees and R* -tree approaches will be used to manage database objects.

Due to voluminous nature of geospatial data, whole data cannot be stored at single server. Also various agencies are dealing with different data formats. According to their usage, geospatial data

needs to be outsourced on some trusted servers by applying some modification/transformations. Transformation should be secure and efficient and robust against attack models. Moreover, the time required to perform transformation/inverse transformation should be less so that user can access it instantly. There is a need to develop secure outsourcing approach using efficient transformation mechanism and also a scheme should be developed which will enable authenticated end users to retrieve data of his own interest (for which he is authorized) without exposing other data. Attack models for both of these schemes need to be exploited.

Whenever data is outsourced, security becomes the first concern. The primary requirements of secure data outsourcing are confidentiality of the outsourced data, user as well as data privacy, and correctness of query results. To achieve security, geospatial data should not be outsourced as it is. Instead, it should be first transformed into some domain using transformation techniques and then outsourced to third party server. Whenever geospatial data is outsourced, it should deal with following issues

- Querying transformed geospatial data
- Robustness evaluation of transformation schemes against various attack models
- Checking correctness of the result obtained from service provider
- Computational overhead in transformation, inverse transformation, and data retrieval process

Above issues can be solved by outsourcing spatial data using some spatial transformations with cryptographic techniques to make it robust against attacks. R* indexing and efficient searching mechanism can be utilized to decrease computational overhead and improve response time.

Another problem cited with outsourcing is that if user wants to retrieve the data in particular special window for which he is authorized, what mechanism should be used so that only that much data is communicated to the user without revealing other data. This problem can be solved by providing efficient index structure, fast search methods and good access control mechanism.

Distribution Layer

While dealing with geospatial data at dissemination level, geospatial data needs to be protected; otherwise it could result in illegal copy and distribution. Use of digital watermarking for copyright protection of geospatial data is the best solution available to deal with this situation as it becomes easy to locate legitimate owner (user/server) of the data in case of tampering and forgery. Also, cryptographic algorithms like identity based encryption, authentication, and use of digital signature make it more robust against various attacks.

Watermarking algorithm used for copyright protection of geospatial data can vary according to its characteristics [15]. They are mainly classified as raster and vector data watermarking algorithms. Also, they will be designed by considering their requirements for copyright protection. The requirements for vector data watermarking are:

- Precision should be preserved.
- Positional accuracy should be maintained.
- Topological relationship should be maintained
- Good robustness against attacks should be provided.
- Watermarking scheme should be invisible and blind.

As these techniques are used for security purpose, it is necessary to hide (invisible) watermark so that no one can suspect and purposely try to destroy it. Also, the end users are required to be authorized and watermark should be retrieved without using original geospatial data and/or watermark. In vector data, the parameters like accuracy, topological relationship and precision have their own importance; therefore it is required to preserve them so that it won't give wrong

results when further analyzed. The robustness of the algorithm should be evaluated against common attacks as well as some specialized attacks. Similarly, for raster data, the requirements are:

- Selective: The watermarking technique should not distort certain specific areas in the image.
- It should be "near Lossless" - Pixel modification is accepted if at every pixel, modification is within user defined distance from original pixels. It is said to be near lossless if this distance is low (i.e. close to original unmodified pixels [16].
- Good robustness against attacks should be provided.
- Watermarking scheme should be invisible and blind.

While dealing with security at various levels, it becomes crucial to handle geospatial data without compromising it. Depending on security techniques, some tailored attacks and a general attacks [17] have to be devised and evaluated against sample geospatial data.

4. CONCLUDING REMARKS AND FUTURE RESEARCH

With the fast development of Internet and communication technology, it becomes easy to copy or distribute the geospatial data. Therefore copyright protection, authenticity, privacy, and spatial data source tracing have become important issues. In this paper, we have presented a conceptual framework to fulfill these requirements. The proposed framework not only poses security at database, but also at distribution level by applying various security techniques. Some of the research directions are outlined. Future work comprises development of integrated security techniques which can be used at both at storage and distribution levels and their performance evaluation with the help of attack models specifically designed for geospatial data.

REFERENCES

[1] Bertino, E., Jajodia, S., Samarati, P.: A Flexible Authorization Mechanism for Relational Database Systems. In: ACM Transactions on Information Systems, 17(2), 101-140(1999)

[2] Gertz, M., Jajodia, S. (editors): The Handbook of Database Security. Applications and Trends. Springer(2007)

[3] Bertino, E., Bonatti, P., Ferrari, E. Trbac: A Temporal Role-based Access Control Model. In: ACM Transactions on Information and System Security, 4(3), 191-233, 2001

[4] Sandhu, R., Coyne, E., Feinstein, H., Youman, C.: Role-based Access Control Models. In: IEEE Computer, 29(2), 38-47(1996)

[5] Agarwal, S., Sprick, B., Wortmann, S.: Credential based Access Control for Semantic Web Services. In: AAAI Spring Symposium- Semantic Web Services, 1(1)(2004)

[6] Wright, T.: Geographic Information Systems. Ontario Office of Information and Privacy Commissioner(1997)

[7] Atluri, V., Chun, S.: An Authorization Model for Geospatial Data. In: IEEE Transactions on Dependable and Secure Computing, 1(4), 238-254(2004)

[8] Belussi, A., Bertino, E., Catania, B., Damiani, M., Nucita, A.: An Authorization Model for Geographical Maps. In: 12th ACM International Workshop on Geographic Information Systems, Washington D.C., U.S.A(2004)

[9] Yiu, M., Ghinita, G., Jensen, C., Kalnis, P.: Enabling Search Services on Outsourced Private Spatial Data. The International Journal on Very Large Data Bases, 19(3), 363-384(2010)

[10] Devanbu, P., Gertz, M., Martel, C., Stubblebine, S.: Authentic Data Publication over the Internet. Journal of Computer Security, 11(3), 291-314(2003)

[11] Dorairangaswamy, M.: A Novel Invisible and Blind Watermarking Scheme For Copyright Protection of Digital Images. International Journal of Computer Science and Network Security, 9(4), 71-77(2009)

[12] Huo, X., Seung, T., Jang, B, Lee, S., Kwon, K.: A Watermarking Scheme Using Polyline and Polygon Characteristic of Shapefile. In: proceedings of third International Conference on Intelligent Networks and Intelligent Systems, Shenyang, pp. 649-652(2010)

[13] Wang, X., Huang, D., Zhang, Z.: A DCT-based Blind Watermarking Algorithm for Vector Digital Maps. Journal of Advanced Materials Research, 179(180), 1053-1058(2011)

[14] Hanashima, M.: Consideration for Information Security Issues in Geospatial Information Services of Local Governments. International Association for Social Science information Services and Technology, pp.16-26(2005)

[15] Jianghua, C., Anbo, L., Guonian, L.: Study on Multiple Watermarking Scheme for GIS Vector Data. In Proceedings of Eighteenth International Conference on Geoinformatics, Beijing, China, pp.1-6(2010)

[16] Barni, M., Bartolini, F., Cappellini, V., Magli, E., Olmo, G.: Near-lossless Digital Watermarking for Copyright Protection of Remote Sensing Images. International Geoscience and Remote Sensing Symposium, Toronto, Canada, pp.1447–1449(2002)

[17] Lin, B., Li, A.: Study on Benchmark System for Copyright Marking Algorithms of GIS Vector Data. In Proceedings of Eighteenth International Conference on Geoinformatics, Beijing, China, pp.1-5(2010)

AUTHORS

Sangita Zope Chaudhari received M.E degree in computer engineering from Mumbai University, India. Currently she is a Ph.D student in Centre of Studies in Resources Engineering at Indian Institute of Technology Bombay, India. Her research interests include digital image processing, advanced database and information systems, and information security techniques.

Parvatham Venkatachalam received the M.Sc. and Ph.D. degree in mathematics from Indian Institute of Technology Bombay, Mumbai, India in 1972 and 1978 respectively. Currently, she is a professor in Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, India. Her research interests include development of GIS for natural and human resources applications, digital image processing of remote sensing satellite data, development of spatial decision support systems, data structure in spatial databases, spatial data mining and warehousing.

Dr. Venkatachalam is a founder member of Indian Society of Geomatics, Life Member of Indian Society of Remote Sensing and Life Member of Indian National Cartographic Association.