Data Security and Access Control for Geospatial Database sets using Novel StegoHash Algorithm

Mamta Malik
Research Scholar
DCRUST, Murthal

Dr.A.K.Sharma
Professor & Dean
YMCA University of Science & Technology,
Faridabad

ABSTRACT
The applications of spatial database (2D and 3D vector maps) are more and more popular in the computer and network environments. It is the reason why information security is an important issue. Data hiding schemes may include map data authentication, secret communication for the purposes of copyright protection, integrity authentication, or secret communication. The purpose of this paper is to proposed novel StegoHash algorithm to enhance the security of spatial database. This algorithm is more appropriate for hiding data in vector maps because the distortions can be removed after the hidden data have been extracted. Combining the concepts of traditional cryptography and steganography, a new spatial database cryptosystem is proposed. Here access remains secure in case of steganalytic attacks especially for highly secure areas like defence, research centres etc.

Keywords
Spatial database; Stenography; StegoHash

1. INTRODUCTION
Currently, most of geospatial data, both spatial and non-spatial properties, are managed by geospatial databases. With the growing need of integration and sharing of geospatial data located in different places on the network, distributed geospatial database technology has become a hot research field. The research topics on distributed geospatial database typically include global spatial data catalogue, global spatial indexing, global query processing and optimization, transaction management, *security control* [1], [19], [22] etc. Meanwhile, geospatial information security has attracted more and more attention in practice. It typically contains three aspects: confidentiality, integrity, and access control. Confidentiality and integrity have already had satisfactory solutions in IT domain, such as cryptography and digital signature. Access control is actually to ensure that legal users execute permitted operations on the intended data of the system. However, access control mechanisms directly provided by database can only implement coarse-grained (map layer) access restriction to spatial data, and have not taken their geometric properties into consideration. Therefore, we have to find out appropriate mechanisms, which shall satisfy both fine-grained (spatial object or even smaller) and geometric access control requirements for spatial database.

Information systems in network environment are confronted with much more threats, therefore, security, especially access control is significant for distributed geospatial databases. There are two major challenges to realize access control for distributed geospatial database. The first is how to define fine-grained control granularities. The general method is to modify table schemas to store authorization information for individual records and fields. The second is how to check access requests against authorization information and make access decision. We propose using views to implement access control functionality, which exploits the built-in mechanisms of distributed geospatial databases to realize fine-grained and spatial access control. It has the following advantages: easy implementation, low overhead and fair flexibility. First of all, we discuss spatial databases in detail where we have to apply security algorithm. Secondly how stenography is helpful in data access. Finally, we proposed the novel algorithm through which procedures of access control in geospatial databases is explained. After the user application with the assigned distributed database account is authenticated by distributed geospatial database, it then gains the account's privileges to the specified view, and data not in views are hidden. As we know, steganography is one of important branches of data-hiding. The purpose of steganography is to send secrete messages under the cover of a carrier signal, i.e., secret communication. On the other hand, steganalysis is the set of techniques that aim to distinguish between cover-objects and stego-objects, or go one step further and estimate some parameters of the embedded message such as its length, location, etc. Several approaches have been proposed to solve the image steganalysis problem and we can broadly classify them into the following groups [14]: Supervised learning based steganalysis [15], Blind identification based steganalysis [16], parametric statistical steganalysis [17] and Hybrid techniques. Each of these methodologies has pros and cons. Therefore, it is up to the user (steganalyst) to choose an appropriate methodology [14]. To our knowledge, there is no work which focuses on steganalysis schemes against StegoHash based steganography methods.

In this paper, we propose a steganalysis method which attacks and successfully identifies the existence of embedding done by the StegoHash data-hiding schemes for 2D and 3D vector maps [4]. Our steganalysis method can even estimate the level of complexity. In section 2 of the paper, the two data-hiding schemes for 2D vector maps [4] are briefly reviewed. Our proposed steganalysis method is presented in section 3. In section 4, simulation results from the application of the proposed steganalysis are presented. Conclusion of the paper is found in section 5.

Steganography and spatial databases
Nowadays, applications of 2D and 3D vector maps [1], [19] have been increasing rapidly. For the purposes of copyright

protection, integrity authentication, or secret communication, etc., the technique of data hiding has been introduced into n-D vector maps. However, due to the strict application requirements of vector maps, modifications to map data are generally undesired. Therefore, reversible schemes are more appropriate for hiding data in vector maps because the distortions can be removed after the hidden data have been extracted. Despite the fact that quite a few data hiding algorithms for images have been proposed, few works have focused on the reversible data hiding algorithms for 2D vector maps. M. Voigt et al. [12] first proposed the method of reversibly hiding data in vector maps. They hide the data by modifying the integer discrete cosine transform (DCT) coefficients of the map coordinates. Steganography means hiding data within media, such that the very existence of data is hidden. It varies from encryption in the sense that its objective is secret communication, rather than data protection. The goal of steganography is to avoid drawing suspicion to the transmission of a hidden message. The plain medium used for hiding information is called cover medium, and the medium generated after embedding the secret text is called stego medium. The use of steganography is thousands of years old; however steganography using digital media is a recent application. Today, computer and network technologies provide easy-to-use communication channels [20] for steganography [2], [3], [7], [9].Essentially, the information-hiding process in a steganography system starts by identifying a cover medium's redundant bits.

The embedding process creates a stego medium by replacing these redundant bits with data from the hidden message. Algorithms in this domain include the LSB (least significant bit) technique, DCT technique and their variants. We discuss the implementation of various steganography algorithms presently being employed today and our novel StegoHash algorithm to ensure compatibility between cover image and hidden data. We then discuss the use of hashing for authentication, to ensure the integrity of secret data access over spatial datasets.

The advances in Remote Sensing (RS), Geographic Information System (GIS), and other survey technologies [5], [6] over the last two decades have dramatically increased geographic information resources in diversity as well as in size. And many organizations have their own database and spatial data with diverse formats. On one hand, interoperability of these existing datasets would integrate these spatial data for better services; on the other hand, different organizations would like to control their own spatial resources and only service for authorized users.

2. REVIEW OF STEGNOGRAPHIC TECHNIQUE

The goal of steganography is to hide messages inside other harmless messages in a way that does not allow any enemy to even detect that there is a second secret message present [8], [11] and shown in figure 1.

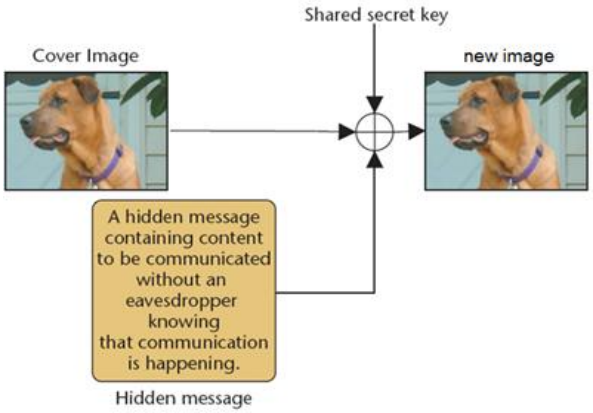


Figure 1 Concept of Steganography

2.1 Masking & Filtering

Information is hidden inside of a image using digital watermarks [3], [7], [13] that include information such as copyright, ownership, or licenses. Algorithm & Transformation are the key component here. This technique hides data in mathematical functions that are often used in compression algorithms.

2.2 Least Significant Bit Insertion (LSB)

This is most common and popular method of modern day steganography. Thus the overall image distortion is kept to a minimum while the message is spaced out over the pixels in the images. This technique works best when the image file is larger than the message file and if the image is grayscale. We replace the LSB [10] of each byte with our secret data described in figure 2.

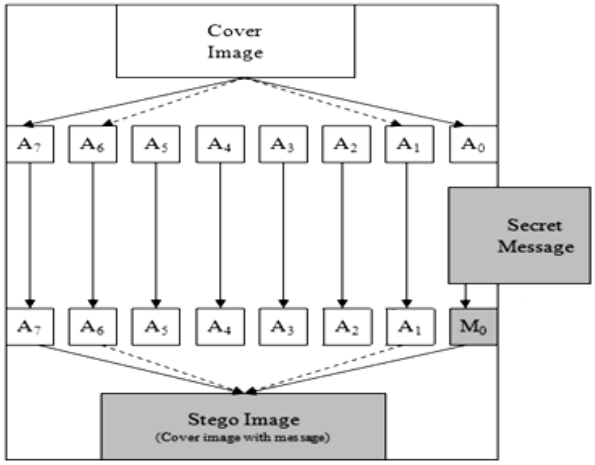


Figure 2 LSB value is replaced by Secret Message

2.3 Adaptive Steganography (based on textures)

What is texture?

- Texture can be defined as a repetition of patterns along a region. These patterns are formed by elements with specific features like size, form, color and direction.

Adeptive steganography?

- Adaptive Steganography [6] considers the features of the cover medium to identify the best region to hide the data. As functioning is shown in figure 3.
- A good place to hide data in digital images is a region with high contrast, several textures and many variations in its pixel levels.

It is so because those regions generally are very noisy, and noise added for hiding a message is difficult to detect**.**

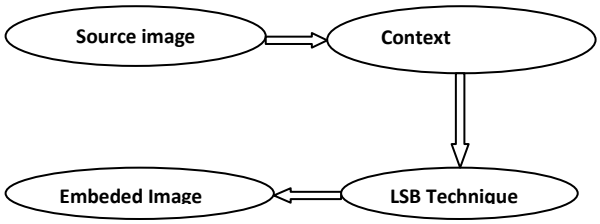


Figure 3 Functioning of Adeptive steganography

Context Algorithm

- Used to identify regions with non-homogeneous textures.
- From those regions we select some pixels to embed a message. Then we apply the LSB technique to hide information.
- The pixel selection process is the following:

Divide the image in non overlapping block of size 3x3 pixels as in figure 4

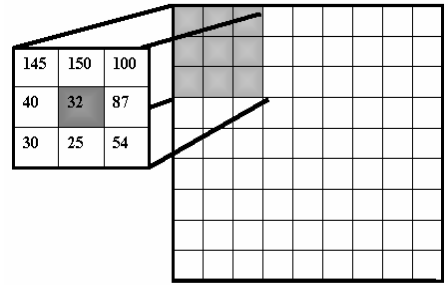


Figure 4 3x3pixel block size of image

Divide each block in four sub-blocks as in figure 5.

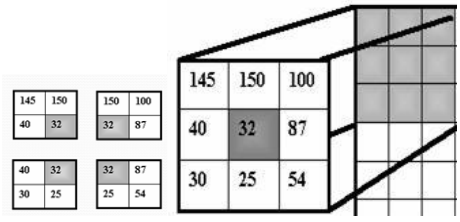


Figure 5 (a) 2x2 block size of 3x3 block size (b) of image

Each sub-block is good if there are at least three different pixel levels. Select the central pixel if the four sub-blocks are good before and after embedding.

2.4 Hashing for Message Authentication

Hash functions are frequently called message digest functions. Their purpose is to extract a fixed-length bit string from a message (image, documents, etc.). Hash functions [3], [21], [23] have found varied applications in various cryptographic, compiler and database search applications. In cryptography, hash functions are typically used for digital signatures to authenticate the message being sent so that the recipient can verify its source. Recently, there has been interest in using hash functions in multimedia applications both for security and indexing. A key feature of conventional cryptographic hashing algorithms such as message digest 5 (MD5) and secure hash algorithm 1 (SHA-1) is that they are extremely sensitive to the message, i.e. changing even one bit of the input will change the output dramatically. However, multimedia data such as digital images go through various manipulations such as compression, enhancement, cropping, and scaling. An image hash function should instead take into account the changes in the visual domain and produce hash values based on the image's visual appearance. Such a hash function would be useful in identifying images in databases, in which the image possibly undergoes incidental changes (such as compression and format changes, common signal processing operations, scanning or watermarking).

A second significant application of a perceptual image hash could be for robust image authentication. In such cases, the hash must be invariant under perceptually insignificant modifications to the image but detect malicious tampering of image data. Several other applications can be identified in the areas of watermarking and information embedding in images. We propose its use as an authentication mechanism coupled with steganography [10], [18]. The hash value of stego image can be calculated and transmitted with the image. Any tampering caused by an adversary in an attempt to foil steganography will be easily detected by the receiver who can discard any such stego image that shows anomaly.

3. PROPOSED ALGORITHM

StegoHash (Stego#), a proposed algorithm is very efficient comparatively with existing steganography techniques in various aspects as we described in table below. There are only few algorithms for security of spatial datasets and their access control. This algorithm is emphasis on the data security and data view privileges at different level of peoples especially for highly sensitive areas like research centers and defence etc. In this algorithm, we need an escape ($) character to separate the message and hash value (#). Therefore, we have to follow all the rules to follow escape character similar to those used in Unix (Like if we have to use that escape character in the message, then write it two times.)

3.1 Algorithm at Sending End

Step 1 Compute the hash (#) value of the secret message by taking the hash value 1000 times.

Step 2 Append escape character '$' to the secret message.

Message=message+"$"

'$'&&temp [i+1]! ='$'

Step 3 After appending the escape ($) character to the secret message, append the hash (#) value. We have got a new message.

Step 4 Hide the new information or message inside the Digital carrier like Image or Audio (called Stego Medium) using any Steganography Algorithm like LSB, DCT etc.

Step 5 Send the Stego Medium to the receiver.

3.2 Algorithm at User Receiving End

Step 1 Extract the data from Stego Medium.

Step 2 Scan the data from left to right.

Step 3 Check for the position of escape ($) character which is not preceded or succeeded by escape ($) character.

Step 4 Compute the hash (#) value of data before the escape ($) character by using the same algorithm that was used on the server side.

Step 5 Check if the computed hash is equal to the hash value stored in the message after escape sequence.

Step 6 If both are equal, hence the Information or message has been intercepted.

3.3 Implementation of StegoHash Algorithm

3.1.1 Algorithm at Sending End

Adding Hash (#) function to information hide in spatial images.

```
addHash (String message)
        {message1=message;
        message=message+"$";}
```

Checking Hash (#) values

```
checkHash(String mess)
        {int flag=0,len=mess.length();
        temp[]=mess.toCharArray();
for(int i=0; i<len-1;i++){
        if(temp[i]=='$'&&temp[i+1]!='$')
        {flag=i+1;}}
                if(flag==0)
        {flag=len-1;}
message=mess.substring(0, flag-1);
givenHash= mess.substring(flag, len);
```

Getting Hash (#) values from the new message

```
temp1=getHash(message1);
        //message1=temp.toString();
        { finall= message+temp1; }
```

Extract data from stego medium at receiver end

```
// Algorithm at User receiving End
newHash=getHash(message);
        if(newHash.equals(givenHash))
                {return true;}
newMessage(String mess)
        {int flag=0,len=mess.length();
        char temp[]=mess.toCharArray();
        for(int i=0; i<len-1;i++)
{if(temp[i]=='$'&&temp[i+1]!='$')
        {flag=i+1;}}
        if(flag==0)
        {flag=len-1;}
Database d=new Database();
                String input1;
                        MessageDigest digest=
MessageDigest.getInstance("SHA-1");
         digest.reset();
byte[] input = digest.digest(mess.getBytes("UTF-
8"));
for (int i = 0; i < 1000; i++)
         { digest.reset();
          input = digest.digest(input);}
 //System.out.println("hash byte: "+input);
                input1=d.byteToBase64(input);
 //input1= input.toString();
                System.out.println("hash:
"+input1);
                return input1 ;}
```

3.1.2 Algorithm at User receiving End

Extract data from stego medium as shown above. First we Scan the data from left to right then check for the position of escape character which is not preceded or succeeded by escape character. Now we compute the hash value of data before the escape character by using the same algorithm that was used on the server side if the computed hash is equal to the hash value stored in the message after escape sequence only then the Information or message has been intercepted.

3.1.3 Advantages

1. The algorithm works for all type of secret data like text, image audio and video. This is possible because we can compute the hash value of Image, audio and video too.

2. The authenticity of the information or message is maintained.

3. It helps in special high precision privileges, detail view privileges and special data view privileges.

3.1.4 Disadvantages

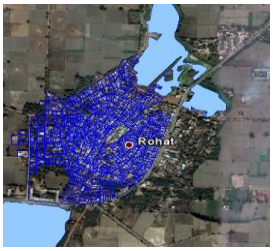
We need to take care of escape character while writing the message because complexity of the algorithm should be maintained.

4. EXPERIMANTAL RESULTS

This algorithm gives better results compared with few other steganography algorithms like LSB in Bmp and GIF, DCT as shown in table 1 by performing experiment on various data. The novel StegoHash algorithm works for all type of secret data like text, image audio and video. This is possible because we can compute the hash value of Image, audio and video too. The authenticity of the information or message is maintained with this algorithm. It helps in special high precision privileges, detail view privileges and special data view privileges.

Table 1 Comparison b/w different steganography techniques with StegoHash

	LSB in Bmp	**LSB in GIF**	**DCT**	**Stego#**
Invisibility	High	Medium	High	High
Payload capacity	High	Medium	Medium	High
Robustness against statistical attacks	Low	Low	Medium	High
Robustness against image manipulation	Low	Low	Medium	High
Independent of file format	Low	Low	Low	Medium
Unsuspicious Files	Low	Low	High	High



(a) **(b)**

Figure 6 (a) is original image where 6(b) is stego image of village Rohat, sonepat, India

Here figure 6 (a) is the original image of Rohat village of district Sonepat, India. After applying novel StegoHash algorithm it will gives result as in figure 6(b). There is no difference between both the images. This algorithm is very efficient to hide a large size file or information without making any noise in original image.

Now this experiment can be used for access control over network especially for spatial databases like Google earth provide an open access to individual. But some areas are highly sensitive with their positional values and information i.e. defence area, research centers, parliaments or government offices of any country etc. Because without security terrorist can use this data for disaster purpose, as what was happen with worlds twin towers in USA. That's why security becomes a very important parameter for every nation security. This is the only purpose to propose concept of access control at different level like high precision level, low precision level etc. For example we can use concept of broadcasting over network where data is available at every user end. But data of different level can be received by only those who are having the privileges to access that data. Using concept of steganography, data is hide within images but can be viewed by those who are having privileges to access detail view of spatial data as things are described below in figures 7. For the purpose there are different snapshoot as describing in figure 7 (a), (b),(c),and (d) are providing access control at different privileges level. (a) Special high precision privileges at house hold level of village Rohat, Sonepat, India (b) Detail view privileges of village Rohat at infrastructure level (c) Detail spatial data/ information view privileges(d) Detail spatial data/ information view privileges uploaded at Google earth.

This is time of information technology where all information is available over the network. But technology should be extended where restriction to access data must be applied. It helps in special high precision privileges, detail view privileges and special data view privileges depends on individual. Where authority can put some constraints on available data based on user's demands decided by assigned authority is essential. For the purpose novel StegoHash () algorithm gives better result and can be extended in future with positive changes by increasing efficiency of algorithm.

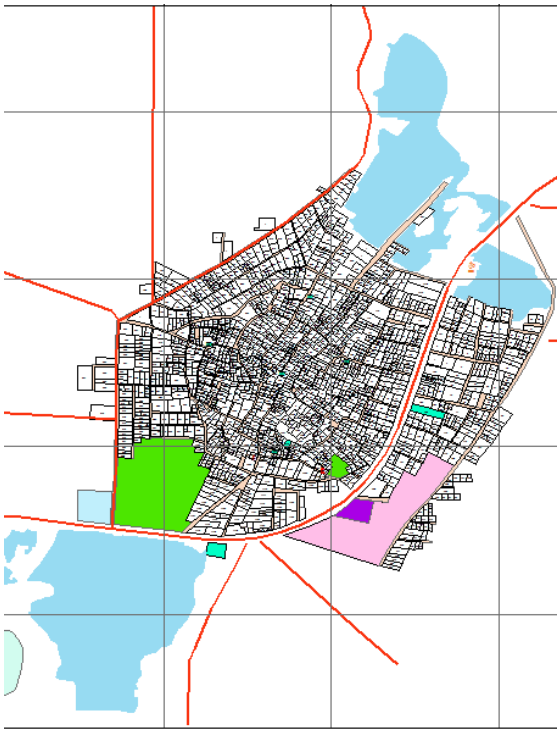


Figure 7(a) Special high precision privileges at house hold level of village Rohat, Sonepat, India

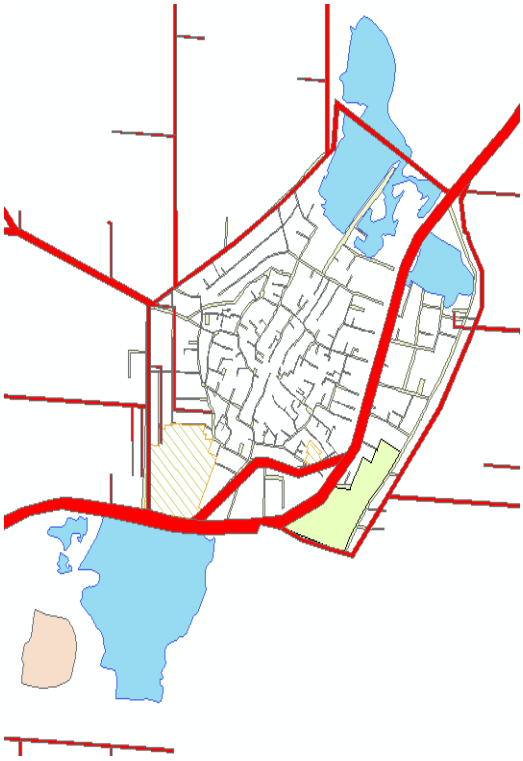


Figure 7(b) Detail view privileges of village Rohat at infrastructure level e.g. Road, Streets, water bodies etc.

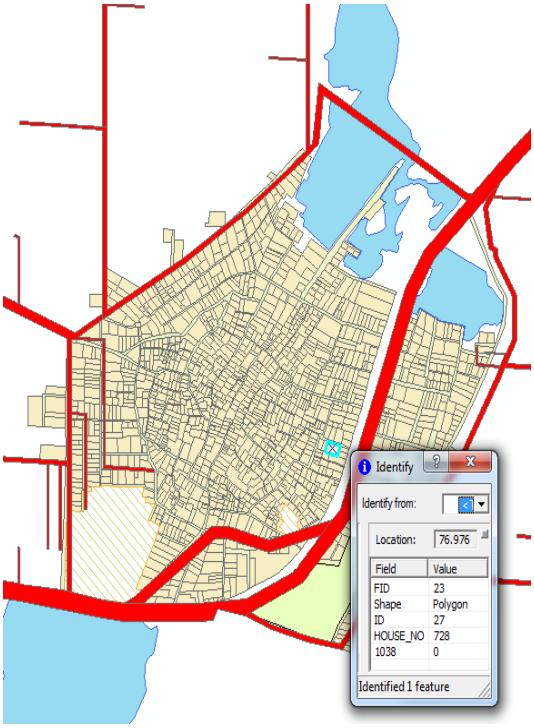


Figure 7 (c) Detail spatial data/ information view privileges

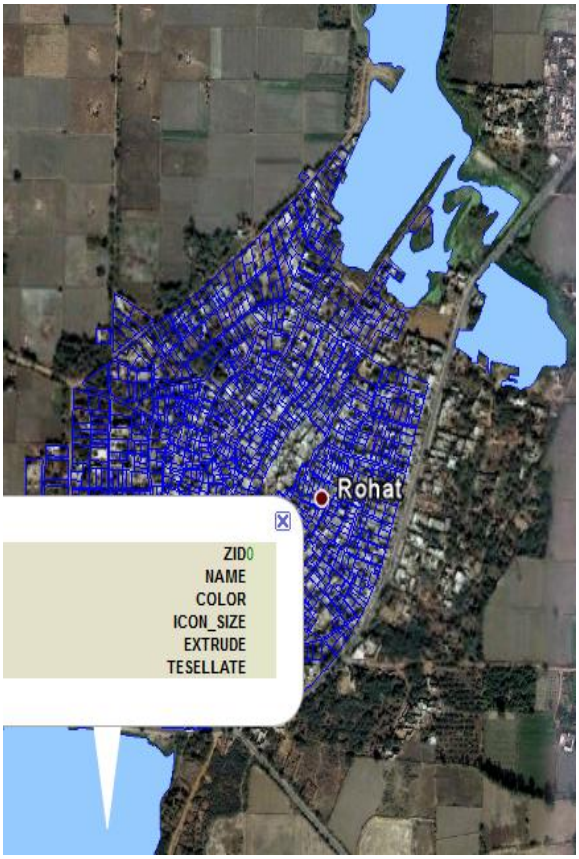


Figure 7(d) Detail spatial data/ information view privileges uploaded at Google earth

5. CONCLUSION

This paper focuses on the steganalysis using StegoHash scheme against other data hiding schemes for 2D and 3D vector maps based on difference expansion. This paper's scheme is effective not only to reveal the presence of secret data, but also to provide access control at different privileges level. The following conclusions can be drawn from the above theory analysis and computing results: 1) Maps having more usable result with high steganalysis accuracy; 2) For different stego-maps, we must choose different access level according to privileges 3) The idea of the proposed schemes is applicable to the 2D and 3D vector maps represented. Moreover, it is possible to extend the scheme to some other data sets, e.g., 3D polygonal meshes, or images in future.

6. ACKNOWLEDGMENTS

The authors would like to thank the YMCA University of Science and Technology and DCRUST Sonepat to support us in our research and allow us to use laboratory for various software and allow accessing various research reports.

7. REFERENCES

[1] R. Ohbuchi, H. Ueda, and S. Endoh, "Robust watermarking of vector digital maps", in Proc. IEEE Int. Conf. Multimedia and Expo, Lausanne, Switzerland, vol. 1,Aug. 26–29, pp. 577–580, 2002.

[2] H. Gou and M. Wu, "Data hiding in curves with applications to map fingerprinting", IEEE Trans. Signal Process., vol. 53, no. 4, pp. 3988–4005,2005.

[3] G. Schulz and M. Voigt, "A high capacity watermarking system for digital maps", in Proc. ACM Int. Workshop on Multimedia and Security, Magdeburg, Germany, pp. 180–186,2004.

[4] XiaoTong Wang, ChengYong Shao, XiaoGang Xu and XiaMu Niu, "Reversible Data-Hiding Scheme for 2-D Vector Maps Based on Difference Expansion", IEEE Transactions on Information Forensics and Security, vol.2,no. 3, pp.311–320,2007.

[5] Stuti Bazaj, Sachin Modi, Anand Mohan, S. P. Singh, "An Improved Algorithm for Data Hiding Using HH-subband Haar Wavelet Coefficients", IJACT, Vol. 2, No. 2, pp. 109 ~ 116, 2010.

[6] Samira Lagzian, Mohsen Soryani, Mahmood Fathy, "A New Robust Watermarking Scheme Based on RDWT-SVD", IJIIP, Vol. 2, No. 1, pp. 22 ~ 29, 2011.

[7] Hongyuan Li, Guangjie Liu, Yuewei Dai, Zhiquan Wang , "Copyright Protecting Using The Secure Visible Removable Watermarking In JPEG Compression ", JDCTA, Vol. 4, No. 8, pp. 34 ~ 42, 2010.

[8] Yongjian Hu, Heung-Kyu Lee, Kaiying Chen, and Jianwei Li, "Difference Expansion Based Reversible Data Hiding Using Two Embedding Directions", IEEE Transactions on Multimedia ,vol.10, no.8, pp.1500–1512,2008.

[9] J. Tian, "Reversible data embedding using a difference expansion", IEEE Trans. Circuits Syst. Video Technol., vol. 13 , pp. 890–896,2003.

[10] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Lossless generalized-LSB data embedding", IEEE Trans. Image Process., vol. 12, no.2 , pp. 157–160,2005.

[11]Mohammad Athar Ali , Eran. A. Edirisinghe, "Reversible Watermarking using Differential Expansion on IPCM Macroblocks in H.264/AVC", JNIT, Vol. 2, No. 1, pp. 105 – 116, 2011.

[12] M. Voigt, B. Yang, and C. Busch, "Reversible watermarking of 2d-vector data", in Proc. ACM Int. Workshop on Multimedia and Security, Magdeburg, Germany, pp. 160–165,2004.

[13] D. M. Thodi and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking", IEEE Trans. Image Process., vol. 16, no.3, pp. 721–730,2007.

[14] Chandramouli R,Subbalakshmi K P, "Current trends in steganalysis:a critical survey", In Proceeding of Eighth International Conference Control on Automation, Robotics and Vision, KunMing: Elseviser Press, pp.964–967,2004.

[15] I. Avcibas, N. Memon, and B. Sankur, "Steganalysis using image quality metrics", IEEE Trans. On Image Processing, vol. 12, no. 2, pp. 221–229,2003.

[16] R. Chandramouli, "A mathematical framework for active steganalysis" ,ACM Multimedia Systems, vol. 9, no.3 , pp. 303–311,2003.

[17] X. Zhang and S. Wang, "Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security", Pattern Recognition Letters 25, pp.331–339,2004.

[18] A.M.Alattar, "Reversible watermark using the difference expansion of a generalized integer transform", IEEE Transactions on Image Processing, vol. 13, no.8 , pp.1147–1156,2004.

[19] J. Lin; Y. Fang, B. Chen, and P. Wu: Analysis of Access Control Mechanisms for Spatial Database. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B8, p. 1443-1448, 2008.

[20] S. Castano, M. G. Fugini, and P. Samarati: Database Security. Addison-Wesley Publishing Company, 1995.

[21] M. Govorov, Y. Khmelevsky, V. Ustimenko, and A. Khorev: Security for GIS N-tier Architecture. Developments in Spatial Data Handling 11th International Symposium on Spatial Data Handling, p. 71-83, 2006.

[22] Z. Yanqun, and W. Qianping: *Security Model for Distributed GIS Spatial Data*. International Symposium on Information Science and Engieering, vol. 2, p.641-645, 2008.

[23] E. Bertino, B. M. Thuraisingham, M. Gertz, and M. L. Damiani: *Security and privacy for geospatial data: concepts and research directions*.SPRINGL 2008.