STUDY ON INTEGRITY OF TOPOLOGICAL RELATIONS IN WATERMARKED GEOSPATIAL VECTOR DATA

Sangita Chaudhari[1], Parvatham Venkatachalam[2], Krishna Mohan Buddhiraju[3]
[1]Professor, ACPCE, Kharghar, Navi Mumbai, India, sschaudhari@acpce.ac.in
[2]Retired Professor, CSRE, IIT Bombay, Mumbai, India, pvenk@csre.iitb.ac.in
[3] Professor, CSRE, IIT Bombay, Mumbai, India, bkmohan@csre.iitb.ac.in

KEYWORDS: Topology, Vector data, Watermarking, Integrity

ABSTRACT

From many years, digital watermarking has been used for copyright protection, data authentication and data source tracing of digital multimedia data. Also, there is significant research on copyright protection of geospatial vector data in recent years. As geospatial data is different from digital images, image watermarking evaluation methods cannot be directly applied to watermarked geospatial data. At present, watermarking algorithms are mainly focusing on robustness evaluation and error analysis. One of the important aspects related to vector data quality i.e. topological relationship integrity is neglected. In this paper, an attempt has been made to incorporate invisible watermark in geospatial vector data by applying multilevel wavelet based watermarking algorithm using different wavelets. The resulted watermarked data has been evaluated in terms of polygon closure and topological relationship integrity. The study helps to choose suitable watermark embedding strength, level of decomposition and wavelet to retain the integrity of topological relationship.

1. INTRODUCTION

The rapid growth of distributed network and Internet provides an easy way to exchange/share digital data among users over a large network. Therefore, it becomes crucial to protect the copyright of the digital data. Digital watermarking is one of the best possible solutions to deal with such issue. Digital watermarking can be applied on different type of data like digital image, audio, video, CAD model, software, and 2D geospatial vector data. Most of the work is done for images, audio, and video but little less attention is paid towards 2D geospatial vector data copyright protection.

Geospatial data is different than other digital data due to its special data structure and application environment. Geospatial vector data is composed of spatial and attribute data. Spatial data indicates the geographical location of real world objects and can be represented as points, polylines, and polygons. All these objects are formed by sequence of coordinates of vertices. Attribute data describes properties of map objects. Information recorded in attribute data is very important and cannot be modified arbitrarily. All the proposed approaches make use of spatial data vertex coordinates for embedding watermark. Specific requirements unique to geospatial vector data need to be considered for vector data watermarking are: precision should be preserved; positional accuracy should be maintained; topological relationship should be preserved; good robustness against attacks should be provided; watermarking scheme should be invisible. Visual quality of the vector map should not get affected by existence of watermark. Also, it should not lead to the element deformation. So it is essential to inspect the qualities of watermarked geospatial vector data. Generally the evaluation is restricted to error analysis like average absolute difference, signal to noise ratio, means square error, and correlation quantity (Cao et al.,2010; Liang et al., 2011; Wang et al.,2009; Yan et al., 2011; Zhang et al., 2010; Zhu et al., 2008).

Another evaluation standard used for evaluation of watermarked data is robustness against attacks. Some of the researchers have used attacks like data conversion, format conversion, data compression, data generalization and many other malicious operations to check robustness of their algorithms. But the relevant research on important aspect indicating data integrity of watermarked geospatial data in terms of topology and shape distortion of vector data features is not considered at all. Therefore, in this paper we have used a watermarking algorithm to embed invisible watermark in geospatial vector data using wavelet based watermarking algorithm (Zope-Chaudhari and Venkatachalam, 2012). In this paper, an attempt has been made to evaluate watermarking algorithm in terms of polygon closure and topological relationship.

2. WATERMARKING OF GEOSPATIAL VECTOR DATA

As wavelet transform reflects the local features better and it is not sensitive to local modification, it is good choice to use wavelet for watermarking. Also, degradation caused using wavelet is less than Fourier based methods and hence better topological preserverance is expected. The polyline/polygon vertex coordinates are taken as an input and one dimensional wavelet transform is applied on them. Low frequency coefficients are modified by embedding the binary watermark using embedding strength P. The choice of wavelet decomposition and wavelets is a crucial issue in wavelet based watermarking. Results vary from application to application for different wavelets as well as wavelet decomposition levels. There are many kinds of wavelets. Depending on application, one can choose wavelet with filters or without filters or wavelets with simple mathematical expression or compactly supported wavelets. One of the simplest wavelets is Haar wavelet. Haar wavelet is discontinuous and resembles a step function. Daubechies family is compactly supported orthogonal wavelet with six coefficients and orthogonality. Similarly, symlet and coiflet are orthogonal compactly supported wavelets with least symmetry and highest number of vanishing moments for a given support width. Biorthogonal and reverse biorthogonal wavelets are spline wavelets. Symmetry and exact reconstruction is possible with FIR filters for these wavelets (Jarrard et al., 2001).

3. EVALUATION OF GEOSPATIAL VECTOR DATA

The watermark (copyright) embedded in geospatial vector data have impact on the geospatial data availability. If watermark size is too large, more geometric errors get generated and topological (spatial) relationship between the features gets destroyed. But according to geospatial vector data characteristics, there should not be significant loss in data accuracy, reduction in data quality, and visual observations. Therefore, while evaluating watermarking algorithm along with error analysis and robustness, spatial relations consistency checking is also important.

3.1 CLOSURE INSPECTION

Geospatial data is made up of features like points, lines, and polygons. Point is the basic logical unit and represented using x-y coordinates. Line is made up of set of coordinate points. Closed areas which are modeled as polygons can be represented as a set of lines with the last point co-ordinates to be the first point coordinates. In embedding process, redundant data gets modified. The modification depends on the length of watermark, coordinates selection to embed watermark and the strength of watermarking. It may result in disclosure of the map objects/features involved in watermarking. Therefore, it is necessary to check for the closure of the features (polygon). The coordinates of first point(X_1, Y_1) and last point (X_n, Y_n) of polygon feature are compared. If they are same, means polygon is closed and watermarking algorithm does not affect the closeness property of polygon features.

3.2 INTEGRITY OF TOPOLOGICAL RELATIONSHIP

Spatial topology defines the spatial relationship among map/spatial features based on the primitive relationships of adjacency, connectivity and containment. Semantic interpretation of feature relationships in geospatial vector data can be described through topological rules. Topology is very important in vector data processing and analysis as it clearly reflects the logical relationship between the spatial objects. There is no effect of projection transformation on topological relationship. Also, it is commonly used for querying spatial features and rebuilding geographical entities. As change in feature coordinates due to watermark embedding cause changes in topological relationship, it becomes urgent to study effect of watermarking on topological relationship. Spatial topology can be expressed as equation 1.

$$ST = \sum (F,\ R,\ \ c,\ \ r, M) \tag{1}$$

where F is set of feature classes; R is set of associated topological rules; c is cluster tolerance; r is ranks defined for the participating features; M is topological metadata. The feature classes participating in topologies can be point, line, polygon, and annotations. Rank is used to preserve the coordinates with the highest accuracy. If the distance between features is less than set x-y tolerance, the lower rank features get snapped to the highest rank feature during processing. Cluster tolerance is the minimum tolerated distance between vertices in a topology. During topology validation, all the vertices that fall within the set cluster tolerance are snapped together (ArcGIS, 2003). If two lines end beside each other within a predetermined distance (cluster tolerance) they will be considered the same when validating topology and snapped together to form a single line.

Topological errors are those errors in spatial data which violate the topological relationships or integrity rules which are inherent in the actual data and are lost during the watermarking process. Topological rules are specified using the feature classes like points, lines, and polygons. Some of the commonly specified rules for single line and polygon feature classes are listed in table I.

4. EXPERIMENTAL RESULTS

To evaluate the scheme, various polyline and polygon vector maps have been used. In this experiment, we have evaluated results using eight different wavelets and three decomposition levels. The wavelets we have used are: the Haar wavelet, the orthogonal Daubechies wavelet (db 4), two biorthogonal wavelets (bior 2.2, bior 4.4), two reverse biorthogonal wavelet (rbio3.3, rbio 5.5), the symlet 6 coefficient wavelet and 5th order coiflet wavelet. The watermark of size 20x40 pixels is used for watermarking. Various embedding strength (*P*) on the scale 0 to 1 are considered in watermarking. Increase in embedding strength obviously increases robustness of the watermark, but decreases visual quality of watermarked vector data (Kutter and Petitcolas, 1999). Polyline contour dataset, polygon dataset, polyline road segment data and watermark are shown in figures 1, 2, and 3.

Table 1. Topological Rules for single line and polygon feature class

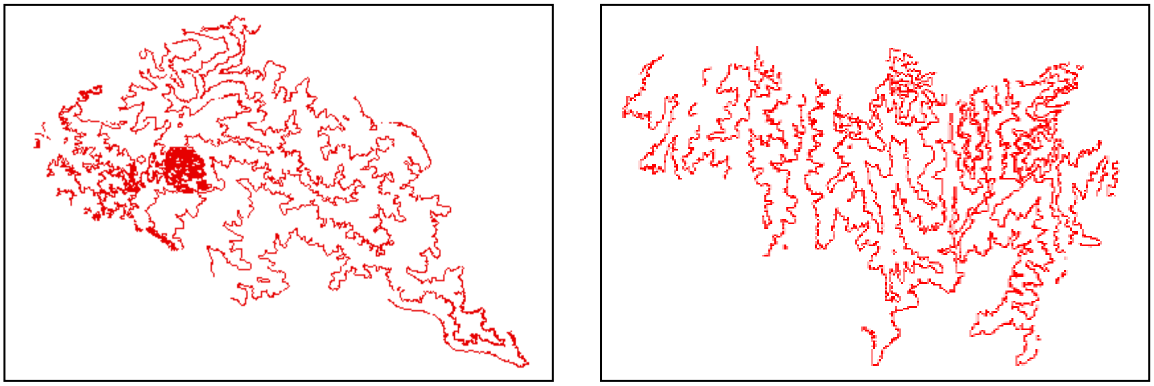
Sr. No.	**Rule Name**	**Feature Classes**	**Description of the Rule**
1	Must Not Overlap	Polygon	Polygons must not overlap within a feature class. Can be dis-connected or touch at a point or along an edge
2	Must not have gaps	Polygon	Polygons must not have voids between them in a feature class
3	Must not have dangles	Line	End of a line must touch any part of one other line or itself. ie , must not have undershoot or overshoot
4	Must not Overlap	Line	Lines must not overlap any part of another line
5	Must not intersect	Line	Line must not cross or overlap any part of another line within the same feature class
6	Must not self-overlap	Line	A single line feature must not overlap itself
7	Must not self-intersect	Line	Lines must not cross or overlap itself. Can touch itself
8	Must be single part	Line	A line feature must have only one part

4.1 POLYGON CLOSURE INSPECTION

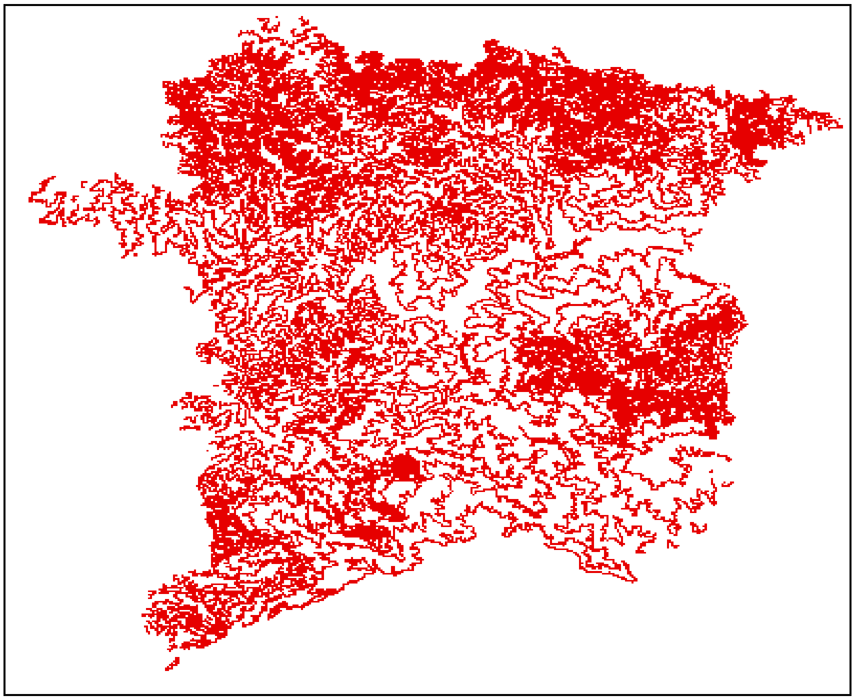
All the polygons of watermarked vector data are checked for closure. Polygon closure is checked by using GIS package GRAM++ (Venkatachalam and Mohan, 2003). Figure 2 shows polygon closure for embedding strength P=0.3, 0.5, and 0.7 respectively at third level wavelet decomposition with Haar wavelet. Red squares indicate the location where polygon is not closed. It is observed that all polygons are closed and number of polygons before and after watermarking is same for watermarking strength *P* between 0 and 0.3. But, polygon starts loosing closeness for $P \geq 0.3$. It clearly shows that as embedding strength increases polygon closure is not preserved.

4.2 TOPOLOGICAL RELATIONS INTEGRITY INSPECTION

Topology relations for single polyline and polygon layer are considered for evaluation. Here, cluster tolerance and rank is considered as 0.001 and 1 respectively for both polyline and polygon data. At each wavelet decomposition level and for each wavelet as well as different embedding strengths, topological relations are checked for watermarked vector data. The topological rules “Must not overlap”, “Must not intersect”, “Must not self-overlap”, and “Must not self-intersect” are used for contour polyline vector data. It has been observed that no lines are overlapping themselves and other lines but some of the lines are crossing themselves and each other. Table 2, 3, and 4 shows the effect of embedding strength on topological relationship for polyline data (polyline Layer1) at each wavelet decomposition level using different wavelets. It has been observed that topological error decreases as decomposition level increases. Also, the embedding strength considered in watermark embedding has enormous effect on topological relationship between the features. Among all wavelets, Haar is giving good results at all wavelet decomposition levels compared to other wavelets. Topology relation inspection is done for all input vector data at third wavelet decomposition level using Haar wavelet for different embedding strength and topological error are shown in table 5. Figure 5 shows topological errors in polyline and polygon vector data at third wavelet decomposition level with embedding strength 0.01.

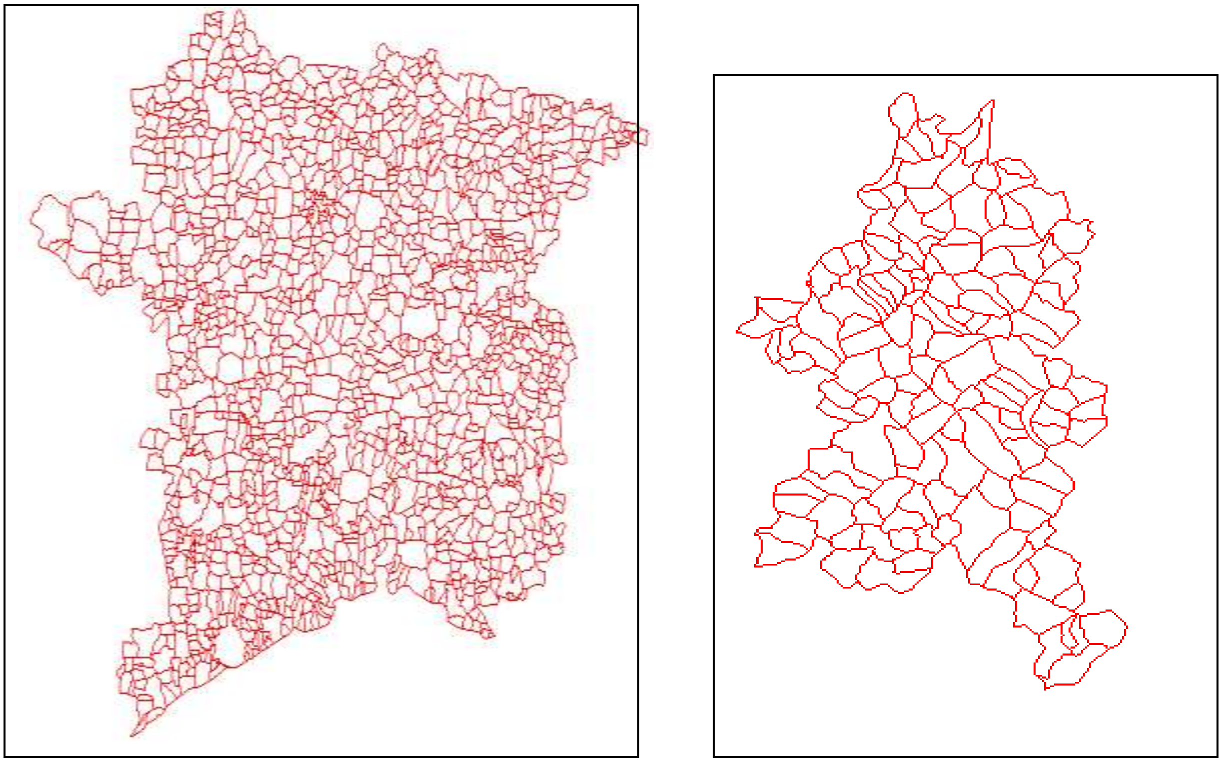


(a) PolylineLayer1 (b) Polyline Layer2



(c) Polyline Layer3

Figure 1. Original Vector Polyline Data



(a) Polygon Layer1 (b) Polygon 2

Figure 2. Original Vector Polygon Data

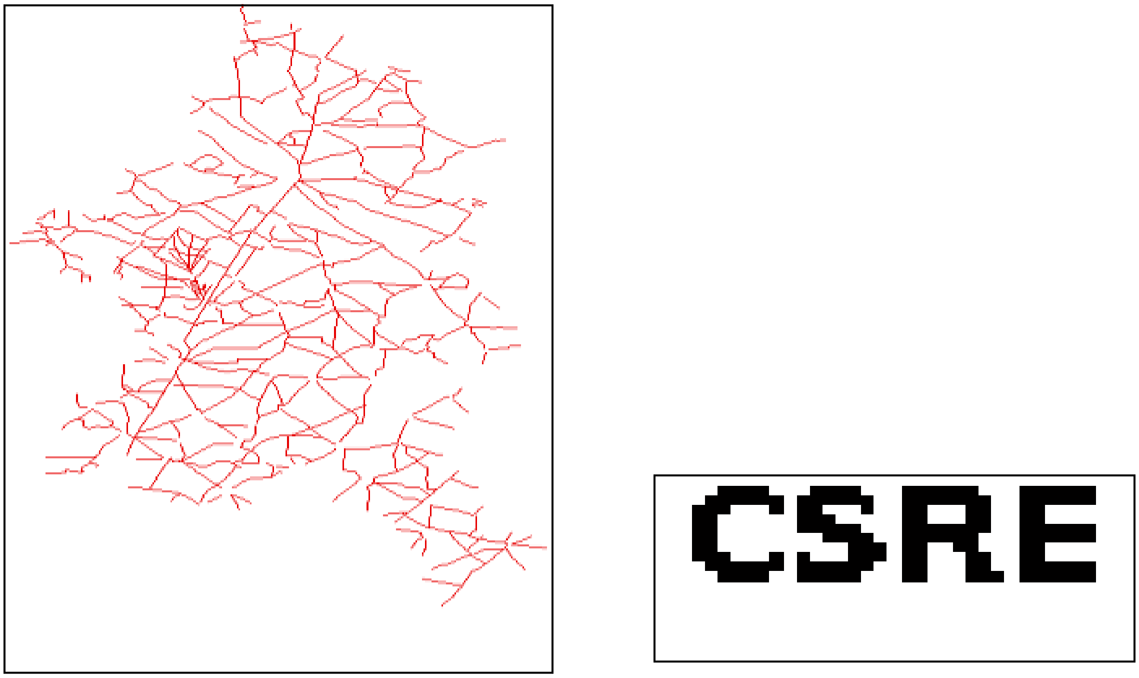


Figure 3. (a)Original Polyline Data (Road_Segment1) and (b) watermark

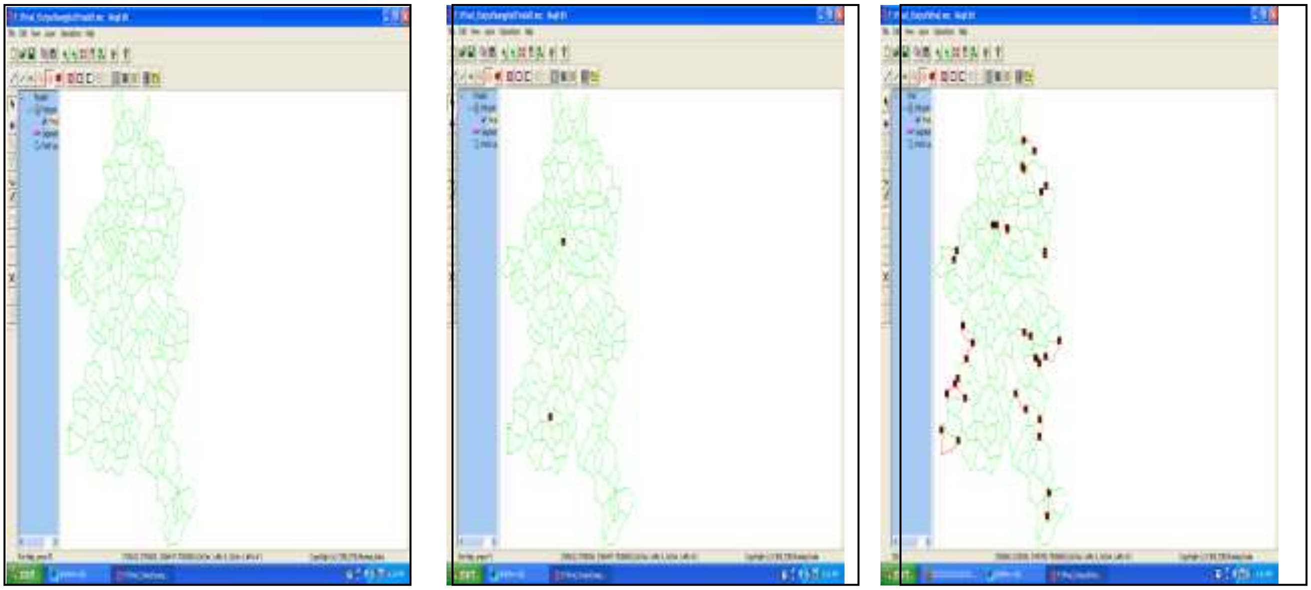


Figure 4. Polygon closure inspection of watermarked vector data (*P*=0.3 and *P*=0.5 and *P*=0.7)

Table 2 Topological Errors at First Level Wavelet Decomposition

Topology Error		**Haar**	**Db4**	**Sym6**	**Coif5**	**Bior2.2**	**Bior3.3**	**Rev. Bior3.3**	**Rev. Bior5.5**
Must be larger than cluster tolerance		2	2	2	3	3	3	3	2
Must not intersect	**P=0.01**	11	12	15	13	12	21	11	12
	P =0.1	11	12	15	13	12	21	11	12
	P =0.3	19	25	27	25	24	34	40	21
	P =0.7	35	41	47	45	37	55	61	56
Must not self-intersect	**P =0.01**	0	0	0	0	0	0	0	0
	P =0.1	0	0	0	0	0	4	0	0
	P =0.3	2	6	5	4	5	6	4	4
	P =0.7	7	11	9	10	11	13	8	9

Table 3 Topological Errors at Second Level Wavelet Decomposition

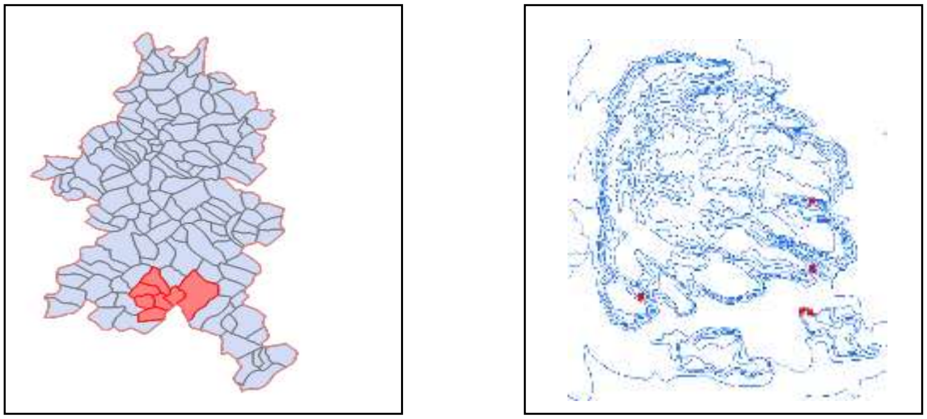
Topology Error		**Haar**	**Db4**	**Sym6**	**Coif5**	**Bior2.2**	**Bior3.3**	**Rev. Bior3.3**	**Rev. Bior5.5**
Must be larger than		0	0	2	0	3	3	3	2
Must not intersect	**P=0.01**	5	7	8	6	7	10	6	7
	P =0.1	5	7	8	6	7	10	6	7
	P =0.3	6	7	9	8	10	13	6	8
	P =0.7	23	28	31	23	30	43	39	27
Must not self-intersect	**P=0.01**	0	0	0	0	0	0	0	0
	P =0.1	0	0	0	0	0	0	0	0
	P =0.3	0	0	0	0	0	3	0	0
	P =0.7	2	5	4	5	3	10	4	3

Table 4 Topological Errors at Third Level Wavelet Decomposition

Topology Error		Haar	Db4	Sym6	Coif5	Bior2.2	Bior3.3	Rev. Bior3.3	Rev. Bior5.5
Must be larger than cluster tolerance		0	0	0	0	0	0	1	0
Must not intersect	P=0.01	3	5	5	3	4	6	4	4
	P =0.1	3	5	5	3	4	6	4	4
	P =0.3	3	10	9	6	9	9	7	5
	P =0.7	6	24	19	8	14	35	28	30
Must not self-intersect	P=0.01	0	0	0	0	0	0	0	0
	P =0.1	0	0	0	0	0	0	0	0
	P =0.3	0	0	0	0	0	1	0	0
	P =0.7	1	2	3	2	4	6	5	4

Table 5 Topological Errors for Different Vector Data at Third Level Wavelet Decomposition

Vector Data	Topology error	Embedding strength(P)			
		0.01	0.1	0.3	0.7
Polyline Layer2	Must not intersect	2	2	4	6
	Must not self- intersect	0	0	0	1
Polyline Layer3	Must not intersect	5	5	9	24
	Must not self- intersect	2	2	5	11
Polygon Layer1	Must not overlap	1	4	4	6
	Must not have gaps	1	1	2	9
Polygon Layer2	Must not overlap	7	7	9	16
	Must not have gaps	1	1	2	9
Road segment Layer1	Must not self- intersect	0	0	2	4
	Must not have dangles	0	0	4	5



(a) (b)

Figure 5 Topological Errors at P=0.01: (a) Polygon data (b) Polyline data

5. CONCLUSION

As geospatial vector data have different and stricter data quality requirements than any other digital data, it is vital to do analysis in terms of closure and topology of watermarked vector data along with imperceptibility, error analysis and robustness evaluation. In this paper, we have evaluated closure of watermarked polygon data and integrity of topological relations for vector watermarked data. Experiments show that although polygon closure is retained for large value of embedding strength, topology starts collapsing even at smaller value of embedding strength of the watermark. Also, visual degradation caused is a function of embedding strength. It has been observed that Haar wavelet outperforms over other wavelets at all decomposition level for given input dataset.

REFERENCES

ArcGIS™(2003) Working with Geodatabase topology an ESRI ® White Paper • May 2003.

Cao, L., Men, C. and Li, X. (2010) Iterative embedding-based reversible watermarking for 2D-vector maps. 17th IEEE International Conference on Image Processing, Hong Kong, 26-29 September, 2010, pp 3685-3688.

Jaffard, S., Meyer, Y. and Ryan, R. (2001) Wavelets Tools for Science and Technology. Society for Industrial and Applied Mathematics (SIAM), 2001.

Kutter, M. and Petitcolas, F. 1999. A fair benchmark for image watermarking systems. In Proceedings of International Conference on Security and Watermarking of Multimedia Contents, Sans Jose, USA, 25-27 January 1999, 226-239.

Liang, B., Rong, J. and Wang, C. (2011) A Vector maps watermarking algorithm based on DCT domain. Journal of Photogrammetry and Remote Sensing, Vol. 38(1), pp 118-121.

Venkatachalam, P. and Mohan, B. (2003) GRAM++ - An indigenous GIS software suite demonstration.
19th International Conference on Data Engineering, Bangalore, India, 5-8 March, 2003, 869-871

Wang, C., Peng, Z., Peng, Y. and Yu, L. (2009) Watermarking 2D vector maps on spatial topology domain. IEEE International Conference on Multimedia Information Networking and Security, Wuhan, China, 18-20 November 2009, pp 71–74.

Yan, H., Li, J. and Wen, H. (2011) A key points-based blind watermarking approach for vector geo-spatial data. Journal of Computers, Environment and Urban Systems 35(6): 485-492.

Zhang, L., Yan, D., Jiang, S. and Shi, T. (2010) New robust watermarking algorithm for vector data. Wuhan University Journal of Natural Sciences 15(5): 403-407.

Zhu, C., Yang, C. and Wang, Q. (2008) A watermarking algorithm for vector geo-spatial data based on integer wavelet transform. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37(B4): 15-18.

Zope-Chaudhari, S. and Venkatachalam, P. (2012) Protecting Geospatial Data using Digital Watermarking. Proceedings International Conference on Computer and Communication Engineering, 3-5 July 2012, Kuala Lumpur, Malaysia, 594-598.