A FRAMEWORK FOR IMPERCEPTIBLITY ANALYSIS OF WATERMARKED DATABASE

Hameed Ullah, Aihab Khan, Basheer Ahmad

Iqra University Islamabd
e-mails: hameedullahpk@gmail.com, aihab@iqraisb.edu.pk, drbasheer@iqraisb.edu.pk
Pakistan

Abstract: In relational database domain, the watermarking schemes presented in literature mainly focus on proof of robustness, whereas, an incomprehensive work is available on proof of imperceptibility of a watermarked database. Most of the work for imperceptibility analysis is in the domain of watermarking multimedia. In this paper, we propose a framework for imperceptibility analysis of watermarked database by using statistical techniques like logistic regression, discriminant analysis and canonical correlation. We also conclude about the usability of watermarked datasets, results in usability increases as imperceptibility increases.

Key words: Box's M test; Canonical Correlation; Discriminant Analysis; Imperceptibility; Logistic Regression; Wald statistics; Wilk's lambda.

1. INTRODUCTION

In order to assure the rights protection, secure and viable mechanisms are essential. Different mechanisms are available and have their own pros and cons. Assurance by legal means is not so effective until this assurance implied by some means of information hiding. Information hiding does not guarantee the actual access control, the question of rights protection remains the same. In order to assure the rights protection, digital watermarking is the suitable tool, which ensures information hiding. The effectiveness of watermarking techniques relies on fact that usability of data is not affected. The resistance against watermark attacks is one of major concerns while designing such techniques.

Digital watermarking techniques can be classified into robust and fragile schemes [1, 2]. Most of the robust techniques introduce errors into the original data which effects data integrity and usability. These techniques can be applied to numeric and categorical data to insert watermarks. Fragile schemes are based on the content features of database relation itself to generate a secure hash, which is stored in least significant bits of original databaseembedding the watermark. Fragile watermarking is aimed to specify any alteration made to the data and is mostly used for integrity authentication [2].

Imperceptibility means that the quality of digital asset is not perceivably distorted and the watermark is invisible to user [3]. It includes how well the watermarked asset under an examination resembles some original asset. Best watermarking schemes assures imperceptibility and overcomes weaknesses like data integrity and data usability, greater distortion will result in low imperceptibility and vice versa.

The existing work on relational database presents watermark insertion and detection algorithms [4-9] along with the robustness of attack analysis. Most of these watermarking technique[2, 5-7] introduce distortion in original database contents that affects imperceptibility. The watermarking schemes that introduce greater distortion will result in low imperceptibility and vice versa. Imperceptibility is an important concern for any watermarking system besides robustness, integrity, incremental updatability and blindness. The imperceptibility can be used to identify the strength of a watermarking system. The watermarking techniques which result in highly imperceptible, can be considered stronger and vice versa.

2. RELATED WORKS

There have been considerable research efforts for imperceptibility analysis in the domain of watermarking multimedia i.e. audio [10-16], video [3, 17-21] and image [2, 22-28]. To the best of our knowledge, a little amount of work found on imperceptibility analysis of watermarked database.

Al-Haj [11] proposed a non-blind digital audio watermarking algorithm and used SNR(signal to noise ratio) for imperceptibility analysis. Both listening tests, i.e. subjective (implemented by human listeners) and objective (implemented by software packages), plays an important role to estimate the imperceptibility of audio watermarking algorithms.

A video watermarking technique presented by Xuemei, Quan [17] is based on shot segmentation and block classification method. Their scheme uses PSNR (Peek signal-to-noise ratio) to measure the imperceptibility of watermarked frames. The video quality assessment methods are also subdivided into two classes: objective and subjective.

Qi [29] proposed two improved perceptual quality methods for imperceptability analysis of watermarked images. One is WPSNR (weighted peak signal to noise ratio) and other method is based on Watson HVS (human visual system) model. Image quality metrics are classified into six classes according to the type of information they use [10]i.e Pixel difference-based, correlation based, edge based, spectral distance based, context based and HVS(human visual system) based.

2.1. Database Watermarking

Agrawal and Kiernan [30]proposed the first well known database watermarking scheme based on small amount of errors. [30] proposed a scheme which selects the tuples, attributes and bit positions to be marked based on an algorithms. The algorithm uses private key K, concatenated with the primary key of the tuple to be seeded for the pseudorandom number generator. Hence, if each value of an attribute is equally likely to be selected and it is as likely that the value will be watermarked then the mean and variance will not be affected significantly. AK also reported the effect of watermarking on the mean and variance of values of marked attributes.

[2] proposed the blind reversible relational database watermarking scheme. The proposed scheme is extension to Agrawal and Kiernan [30]algorithm. This scheme is used to prove the true ownership of the database's owner, and recovers the original database relation once the watermark information is detected and authenticated. For this, a stream of owner-specific watermark bits inserted into the fractional portion of numeric attributes by using a reversible data-embedding technique called prediction-error expansion. When a database relation is outsourced and the owner suspects any intentional ill use, he can go by detection

process to verify whether the outsourced relation is abused by detecting and verifying the embedded owner-specific watermark information. For imperceptibility the variance used as statistical metric to measure the quantitative change introduced to the data.

[31] used non-parametric statistical test to determine difference between original and watermarked database. The Kruskal-Wallis test compared the means difference of each attribute and decided imperceptibility of watermarked database. However, this test could not identify where the difference between original and watermarked database occurred. Three important issues identified in existing work.

- It analyzed differences of individual variables instead of whole database.
- Significance level of similarity or difference was undefined.
- Contribution of individual variable in watermarking dataset was also unknown.

It is evident from previous work there is an immense need of new metrics, which could reflect imperceptibility in a better way. In this article, we propose a framework to assess the imperceptibility of watermarking methods of relational databases. Our work makes possible to give results about the strength of watermarking techniques, which are more logical. The aim is to identify and analyze similarity between original and watermarked datasets for relational databases.

Table 1. Notations used in this paper

Symbol	***Description***	***Symbol***	***Description***
D_0	Original Database	n_j	Number of observations for the j^{th} group
D_ω	Watermarked Database	$\bar{x}_j$	Row vector of the sample mean for the j^{th}group
D	$\begin{bmatrix} D_0 \\ \dots\dots \\ D_\omega \end{bmatrix}$	W_i	Value of discriminant weight of Independent variable i
p	Probability of Watermark	S	Pooled within-group covariance matrix
X_{ik}	Value of the ith independent variable for object k	M	Within group matrix of sums of squares and products
S_j	Within-group covariance matrix for group j	T	Total matrix of sums of squares and cross-products
ξ	Number of least significant bits available for marking in an attribute	τ	Number of most significant bits available for marking in an attribute
a	Intercept	Z_{jk}	Value of the j^{th}discriminant function for the k^{th} case
β_k	Regression coefficient	SE	Standard error
R_0	Correctly classified cases in discriminant analysis	L_0	Correctly classified cases in logistic regression
HR_D	Hit ratio of discriminant analysis	HR_L	Hit ratio of logistic regression
CC	Canonical Correlation	θ	Imperceptibility

3. PROPOSED FRAMEWORK

Figure 1 shows the stages for analyzing imperceptibility on original and watermarked datasets. The notations used in this article provided in Table 1. The watermarked dataset obtained by techniques used by[2, 32]. After, discriminant analysis, logistic regression and canonical correlation analysis used to find out strength of association between original and watermarked datasets. Results combined and interpreted from previous stages based on which conclusions derived. The details of these experiments presented in next section

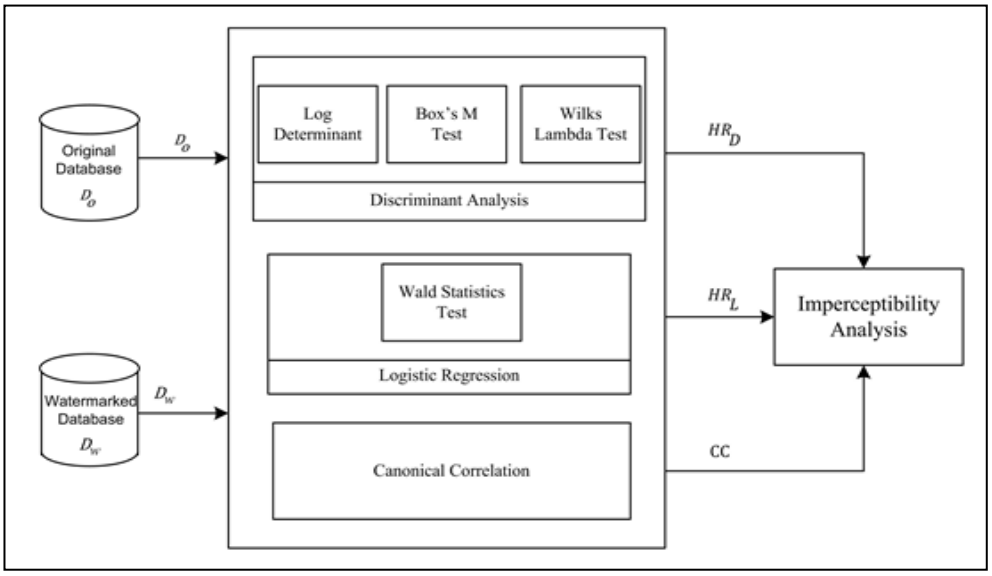


Fig. 1. Watermarking Imperceptibility framework

3.1. Proposed Algorithm

We now present an algorithm for imperceptibility analysis of watermarking database relations.

Input:, D_0,D_ω ,D, R_0, L_0

1. **For** i = 1 to number of databases **do**
2. $C_0 = Cov(D_0)$ //covariance matrix of original database
3. $C_\omega = Cov(D_\omega)$ //covariance matrix of watermarked
4. **For** all attributes **do**
5. $BM = (n - n_j)\, log|S| - \sum_{j=1}^{n_j}(n_j\ 1)\, log|S_j|$ // Box's M test
6. $\Lambda = \frac{|M|}{|T|}$ //Wilk's Lambda Test
7. $HR_D = \frac{R_0}{D}$ // Hit ratio of DA
8. $Z_k = \frac{\beta_k}{SE(\beta_k)}$ //Wald Statistic
9. $HR_L = \frac{L_0}{D}$ //Hit ratio of LR
10. $C_{0\omega} = C'_{\omega 0} = Cov(D_0, D_\omega)$

11. $C_{\omega 0} = C'_{0\omega} = Cov(D_\omega, D_0)$
12. $Z = Cov(D)$
13. $Y = (C_0^{1/2} C_{0\omega} C_\omega^{-1} C_{\omega 0} C_0^{-1/2})$ //By Cauchy-Schwarz inequality
14. $e = eig(Y)$ //Eigen Vector of Y
15. $a_i = C_0^{-1/2} e_i$ //Coefficient vector
16. $b_i = C_\omega^{-1/2} f_i$ //Coefficient vector
17. $U_i = a_i' Z^{(1)}$ //Canonical variates
18. $V_i = b_i' Z^{(2)}$ //Canonical variates
19. $Var(U_i) = a' C_0 a$
20. $Var(V_i) = b' C_\omega b$
21. $Cov(U_i, V_i) = a' C_{0\omega} b$
22. $CC = Corr\ (U_i, V_i) = \frac{Cov\ (U_i, V_i)}{\sqrt{Var(U_i)}\sqrt{Var(V_i)}}$
23. $\theta = \frac{\frac{HR_D + 2HR_L + 3CC}{6}}{\max(\theta)} (0 < \theta < 1)$
24. **End For**

End For

We list five basic inputs used in the algorithm.

1. D_0 = Original database
2. D_ω = Watermarked database
3. D = Original database concatenated with watermarked database
4. R_0 = Correctly classified cases in DA
5. L_0 = Correctly classified cases in LR

At lines 1-3 variance-covariance matrices of original and watermarked datasets are determined, which are then used to calculate logistic regression, discriminant analysis and canonical correlation. Lines 4-7 compute the Box's M, Wilks Lambda and hit ratio of discriminant analysis for each attribute of database. Hit ratio is the percentage of correctly classified cases in a given model [33]. At line 8 Wald statistics is computed which determines the most watermarked attributes. Lines 10-22 determine the canonical correlation. To measure imperceptibility, weights were assinged to variables based on the strength of statistical methods. CC is strongest and assigned largest weight 3, HR_L assigned weight of 2 and HR_D assigned 1 as weight, because logistic regression is robust than discriminant analysis. We compute weighted mean of all variables and divide by maximum value of θ obtained by several experiments in line 23. θ is imperceptibility mertric, ranges between 0 and 1.

4. EXPERIMENTS

This section highlights experimental analysis of imperceptibility analysis on original and watermarked datasets. The experiments were performed on an Intel Core i3, CPU 3.2GHz with 4GB RAM using scientific dataset. Experiments performed by varying the quantity of watermark. When watermark embed in LSB of fractional part of number, the quantity of watermark is low and invisible: it is original case of watermarking algorithm. When watermark embed in MSB of integral part of number, the quantity of watermark is high and visible: it is modified case of watermarking algorithm. The watermarked data sets

are obtained from technique used by [2, 32] using original and modified case of watermark insertion. The original and watermarked data sets are used as inputs for various experiments; such as discriminant analysis, logistic regression and canonical correlation analysis which are presented in this section.

Experiments were performed using the YearPredictionMSD experiment data set, available from the University of California at Irvine KDDArchive. The data set has 515345 rows, each with 90 attributes. We added an extra attribute called id to serve as the primary key and chose the four integer-valued attributes and 70000 rows as candidates for watermarking.

Table 2. Variance difference for [32] algorithm using original case

	γ	10000	1000	100	10	1
Attribute	**Variance**					
TA1	3.06E+06	0	3.63E-01	-3.26E-02	-9.59E-02	-1.71E-02
TA2	1.59E+06	0	7.34E-01	1.52E-01	3.22E-02	-6.93E-01
TA3	1.21E+06	0	2.64E-01	3.60E-02	-2.61E-01	-1.37E-02
TA4	2.25E+05	0	-1.72E-01	-3.29E-02	2.02E-02	-1.19E-01

Table 3. Variance difference for [2] algorithm using original case

	γ	96	48	24	12	6
Attribute	**Variance**					
TA1	3.06E+06	-1.60E-02	-8.14E-03	7.40E-02	1.10E-01	1.54E-01
TA2	1.59E+06	7.27E-04	-5.16E-02	-6.86E-03	5.49E-02	2.87E-01
TA3	1.21E+06	-4.93E-04	-1.19E-02	-3.78E-02	-7.30E-02	4.23E-01
TA4	2.25E+05	8.09E-03	1.26E-02	3.03E-02	3.60E-02	2.60E-01

Table 4. Variance difference for [32] algorithm using modified case

	γ	10000	1000	100	10	1
Attribute	**Variance**					
TA1	3.06E+06	0	2.44E+02	3.16E+03	2.94E+04	5.50E+05
TA2	1.59E+06	0	2.13E+01	2.90E+03	1.45E+04	3.04E+05
TA3	1.21E+06	0	8.48E+01	5.04E+02	8.97E+03	2.32E+05
TA4	2.25E+05	6.00E+00	3.53E+01	1.08E+02	1.24E+03	2.65E+04

Table 5. Variance difference for [2]algorithm using modified case

	γ	96	48	24	12	6
Attribute	**Variance**					
TA1	3.06E+06	4.97E+07	3.30E+07	3.15E+08	9.97E+08	4.63E+09
TA2	1.59E+06	2.90E+07	1.16E+08	2.57E+08	1.33E+09	5.43E+09
TA3	1.21E+06	3.94E+07	1.48E+08	4.44E+08	1.24E+09	5.39E+09
TA4	2.25E+05	5.30E+06	2.25E+08	1.72E+08	1.18E+09	5.56E+09

We next report our results based on statistical metric used by [2, 32] . The results describe the change in variance before and after watermarking. Experiments were performed for (gap between rows) γ =1, 10, 100, 1000 ,10000 and γ = 6,12,24, 48, 96. For original case ξ=8 and for modified τ = 2 were used. We found that for each attribute, variance difference for all original cases was very small however greater changes occured in variance in all modified cases of watermark insertion. Tables 2 and 3 clearly depict that changes are not significant comparative to original values of variance. As expected significant change in variance occured in tables 4 and 5 when τ = 2 becasue of larger pertubation in greater fractions of tuples.

The results presented the quantitative change occured in each attribute after watermarking. Larger variance difference in attributes indicated larger changes, imperceptibility is considered larger and vice versa.

There is an immense need to define the significance level of difference. The above results also did not give a whole picture of the data. We see a scattered picture of the data and decide our own about the change in the data after watermarking. In such situations, discriminant analysis, logistic regression and canonical correlation analysis are the best available methods to compare the original and the watermarked data sets. These methods not only help to compare data with a defined significance level of change but also define a framework which gives a single value θ ranges between 0 and 1 to measure imperceptibility.

4.1. Discriminant Analysis

In this section, experiments performed to determine mean and variance difference for original and watermarked datasets.

4.1.1. Box's M

Box's M tests the hypothesis of equality of covariance over the original and watermarked datasets [34].

$$Box's\ M = (n - n_j)\ log|S| - \sum_{j=1}^{n_j} (n_j\ 1)\ log|S_j| \qquad (1)$$

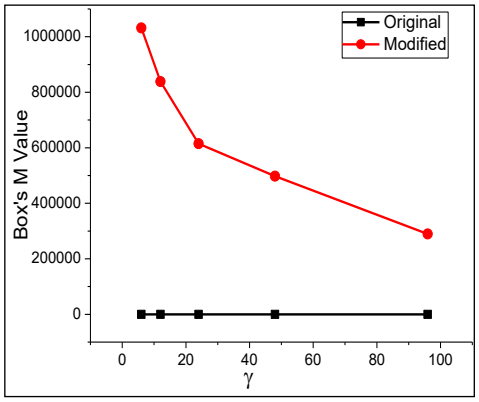


Fig.2. Box's M for [32]algorithm

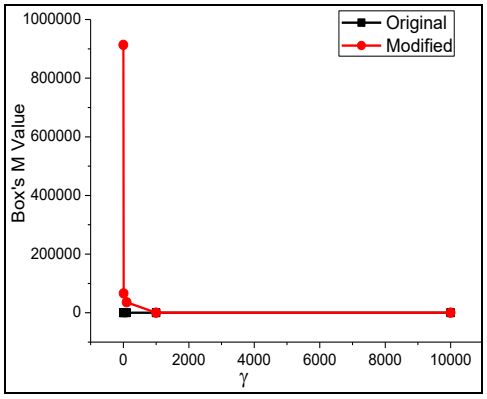


Fig.3. Box's M for [2]algorithm

Figures 2 and 3 demonstrate the change of Box's M test for original and modified watermarked datasets when γ (gap between rows) increased. The trend is that Box's M increased with the decrease in γ. However it followed similar behavior in original case. We saw larger values in modified cases both in 2, 3 while low in original cases. Conclusion drawn from these two figures is that variances of both datasets are statistically same in original case and became statistically different in modified case with the increase in γ. The results of this test does not contribute to proposed framework however this test helped to measure variance difference between origianl and watermarked datasets.

4.1.2. Wilk's lambda

Wilk's lambda tests whether there are differences between the means of identified groups [35]. Wilk's lambda identifies if the difference between the means of original and watermarked database is significant or not.

$$\Lambda = \frac{|M|}{|T|} \qquad (2)$$

Table 6. Wilk's lambda significance for [32] Algorithm Results

γ	1		10		100		1000		10000	
	Original	**Modified**	**Original**	**Modified**	**Original**	**Modified**	**Original**	**Modified**	**Original**	**Modified**
TA1	*.541*	***0***	*.909*	*0*	*.921*	***.001***	*.615*	*.370*	*.274*	*.142*
TA2	*.653*	***0***	*.953*	*0*	*.241*	***.001***	*.954*	*.990*	*.907*	*.098*
TA3	*.806*	***0***	*.166*	*0*	*.240*	***.006***	*.070*	*.546*	*.897*	*.254*
TA4	*.973*	***0***	*.845*	***0***	*.958*	***.013***	*.208*	*.675*	*.212*	*.167*

Table 7. Wilk's lambda significance for [2]Algorithm Results

γ	6		12		24		48		96	
	Original	**Modified**	**Original**	**Modified**	**Original**	**Modified**	**Original**	**Modified**	**Original**	**Modified**
TA1	*.652*	***0***	*.796*	***0***	*.979*	***.001***	*.946*	*.557*	*.285*	*.228*
TA2	*.789*	***0***	*.562*	***0***	*.689*	***.003***	*.894*	*.155*	*.840*	*.273*
TA3	*.542*	***0***	*.749*	***0***	*.746*	***0***	*.451*	*.083*	*.621*	*.400*
TA4	*.898*	***0***	*.866*	***0***	*.754*	***.003***	*.504*	***.008***	*.915*	*.394*

Tables 6 & 7 present the significance value of Wilk's lambda test for first four attributes of database. A low significance value for Wilks lambda test (typically 0.05) demonstrates that there is a significant difference between the means of original and watermarked datasets. Overall, tables 6 & 7, the mean difference between original and watermarked datasets increased with the decrease in γ. When watermark with γ = 1, 6 inserted, the mean difference was very small in original case while maximum in modified case, showing that the difference between the means was significant. When watermark with γ = 48, 96 and γ = 1000, 10000 inserted the mean difference was low and sig value is greater than 0.05 in original and modified case, showing that the difference between the means is not significant. Like Box's

M, Wilk's lambda only tests whether there significant difference between means of original and watermarked datasets. Application of discriminant analysis leads to measure hit-ratio R_0 finally that contribute to proposed framework.

4.2. Logistic Regression

Logistic Regression analyzes the relationship between multiple independent variables and a categorical dependent variable [33].

4.2.1. Wald Statistics

The Wald statistic assess the contribution of individual variable or the significance of individual coefficients in a given model [36]. This test analyzed the attributes with most number of watermarks in the given database.

$$Z = \frac{\beta_k}{SE(\beta_{k)}} \qquad (3)$$

Table 8. Wald Statistics Results for [32] Algorithm

γ	***1***		***10***		***100***		***1000***		***10000***	
	Original	***Modified***	***Original***	***Modified***	***Original***	***Modified***	***Original***	***Modified***	***Original***	***Modified***
TA1	*.565*	***0***	*.751*	***0***	*.455*	*.076*	*.229*	*.228*	*.390*	*.663*
TA2	*.763*	***0***	*.548*	***0***	*.287*	*.087*	*.554*	*.910*	*.469*	*.689*
TA3	*.958*	***0***	*.079*	***0***	*.451*	*.238*	*.490*	*.988*	*.742*	*.347*
TA4	*.624*	***0***	*.778*	***0***	*.581*	*.496*	*.092*	*.400*	*.256*	*.768*

Table 9. Wald Statistics Results for [2]Algorithm

γ	***6***		***12***		***24***		***48***		***96***	
	Original	***Modified***	***Original***	***Modified***	***Original***	***Modified***	***Original***	***Modified***	***Original***	***Modified***
TA1	*.765*	***0***	*.896*	***.004***	*.909*	***.019***	*.666*	*.789*	*.191*	*.463*
TA2	*.389*	***0***	*.492*	***.002***	*.590*	*.062*	*.431*	*.305*	*.898*	*.442*
TA3	*.205*	***0***	*.453*	***.005***	*.825*	***.034***	*.347*	*.274*	*.362*	*.601*
TA4	*.584*	***0***	*.901*	***.034***	*.548*	*.158*	*.392*	*.245*	*.763*	*.513*

Tables 8 & 9 represent the sig values of Wald statistics test. A low significance value for Wald statistics test (typically 0.05) indicates that attribute has most number of watermarks than other attributes in the database or the attribute contribute significantly in watermarking the database. When watermark with $\gamma = 1$ and $\gamma = 6$ inserted, the significance value is greater than .05 in original case and 0 in modified case. This showed that any attribute did not contribute significantly in watermarking the database in original case while the contribution of each variable was significantly high in modified case. When watermark with $\gamma = 24$ inserted the only TA1 had significant contribution in watermarking dataset.

When watermark with $\gamma = 1000$, 10000 and $\gamma = 48$, 96 inserted, contribution of each variable was not significant in watermarking database in both original and modified case. This test helped to identify the most watermarked attributes.

Hit ratio of logistic regreesion L_0 contribute to framework and is obtained by measuring number of correctly classified cases in the given model.

4.3. Canonical Correlation Analysis

Canonical correlation is the correlation between a linear combination of the variables in one dataset and a linear combination of the variables in other dataset [34]. Canonical correlation close to 1 indicates that there is no significant change between original and watermarked datasets.

$$Corr(U,V) = \frac{a'\Sigma_{12}b}{\sqrt{a'C_0a}\sqrt{b'C_\omega b}} \quad (4)$$

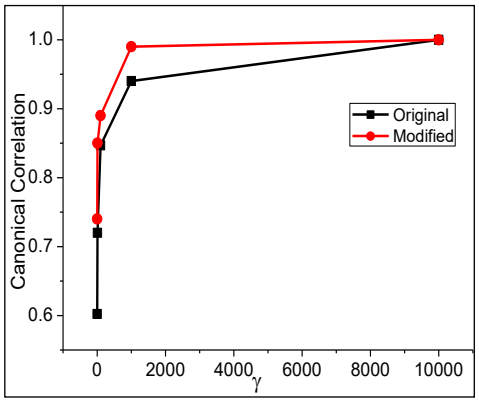


Fig. 4. CC for [32]algorithm

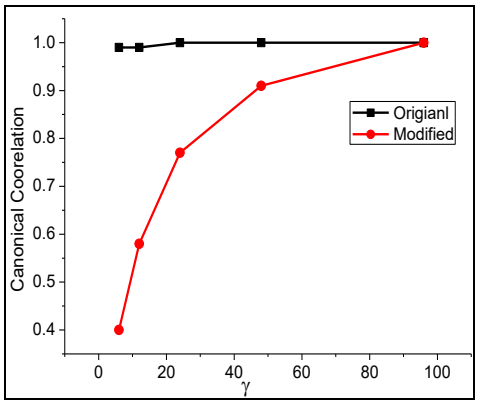


Fig. 5. CC for [2]Algorithm

Canonical correlations between original and watermarked datasets showed in figure 4 & 5. Overall, canonical correlation between original and watermarked datasets increased with the increase in gap between the rows. In Figure 4, when dataset with γ = 1 of watermarks is selected, the canonical correlation was .99 in original case and .68 in modified case, which indicated no significant change occurred after watermark in original case while significant change occurred in modified case.

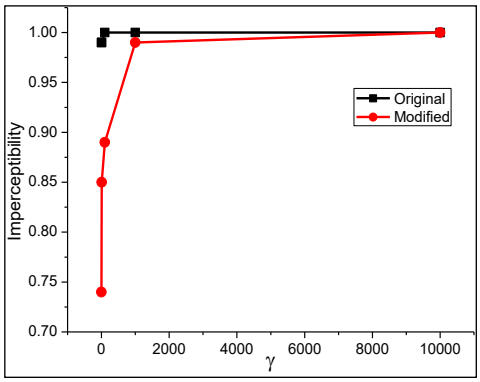
When dataset with γ = 100 of watermarks was selected, the canonical correlation was 0.99 in original case and .84 in modified case, which indicated no significant change occurred after watermark in original case and a significant change in modified case. Canonical correlation was 1 when dataset with γ = 10000 selected in both cases, showed no significant change occurred after watermark in dataset.

In Figure 5, when dataset with γ = 6 of watermarks was selected, the canonical correlation was 0.99 in original case and 0.4 in modified case, clearly indicated no significant change occurred after watermark in original case while significant change occurred in modified case.

Canonical correlation increased with the increase in γ in both cases, when dataset with γ=96 of watermarks was selected the canonical correlation was 1 for both cases which indicated the both datasets are statistically same. These results indicate that a strong relationship exists between original and watermarked datasets for original case while a weak relationship for modified case.

4.4. Imperceptibility

We next report imperceptibility (θ) based on derived framework in figure 6 & 7. Imperceptibility is measured by running both algorithms in original and modified cases. The results presented in graphs show high values of θ in original case, which shows that the original and watermarked data look quite similar and it is difficult for the attacker to identify whether the dataset is original or watermarked. On the other hand θ is lowest in modified case when watermark inserted with $\gamma = 6$ and $\gamma = 1$ and gradually increases with the increase in γ. Both graphs showed datasets were on low imperceptibility in modified case hence quite useful for attacker.



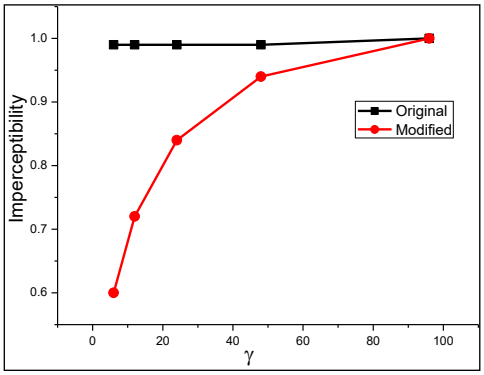


Fig. 6.Imperceptibility for [32] algorithm *Fig.7.Imperceptibility for [2] algorithm*

These results indicate when watermarked datasets are imperceptible, a strong relationship exists between original and watermarked datasets thus the usability of watermarked datasets will be high. When watermarked dataset will be perceptible a weak relationship exists between original and watermarked datasets and it is easy for an attacker to identify that dataset is watermarked. We derived the following relation.

$$\theta \, \alpha \, \frac{1}{\gamma} \qquad (5)$$

5. CONCLUSION

In this article, using three statistical methods a framework proposed to determine the imperceptibility of watermarked dataset. Our framework enables to define significance of difference and imperceptibility. The watermarked dataset used for experiments obtained by two algorithms, though the proposed framework is applicable to any distortion based database-watermarking technique. We conclude that when original and watermarked datasets have similarities as determined by statistical tests so the watermarked dataset has high data integrity and thus data usability is high. Similarly, when original and watermarked datasets are statistically different, the watermarked dataset has low data integrity so the usability becomes low.

REFERENCES

1. Khan, A. and S.A. Husain, A fragile zero watermarking scheme to detect and characterize malicious modifications in database relations. The Scientific World Journal, vol 2013, Article ID 796726, 16 pages, 2013 .

2. Farfoura, M.E., et al., A blind reversible method for watermarking relational databases based on a time-stamping protocol. Expert Systems with Applications, 39(3), 2012, pp. 3185-3196.

3. Abdallah, E.E., A.B. Hamza, and P. Bhattacharya, Video watermarking using wavelet transform and tensor algebra. Signal, Image and Video Processing, 4(2), 2010, pp. 233-245.

4. Zhang, Y., et al. Relational databases watermark technique based on content characteristic. in Innovative Computing, Information and Control, 2006. ICICIC'06. First International Conference on. 2006, IEEE.

5. Ali, Y.H. and B.S. Mahdi, Watermarking for relational database by using threshold generator. Computer Sciences Department, University of Technology Baghdad, 2011.

6. Zhou, X., M. Huang, and Z. Peng. An additive-attack-proof watermarking mechanism for databases' copyrights protection using image. in Proceedings of the 2007 ACM symposium on Applied computing. 2007. ACM.

7. Pournaghshband, V. A new watermarking approach for relational data. in Proceedings of the 46th Annual Southeast Regional Conference on XX. 2008. ACM.

8. Kumar, M., O. Verma, and A. Saxena, Elliptic Curve Cryptography (ECC) based Relational Database Watermarking. International Journal of Computer Applications, 154(6), 2016.

9. Kyriakopoulos, S., T. Tzouramanis, and Y. Manolopoulos. The dbMark: A benchmarking system for watermarking methods for relational databases. in 11th IEEE International Conference on Research Challenges in Information Science (RCIS), 2017.

10. Sayood, K., Statistical evaluation of image quality measures. Journal of electronic imaging, 11(2), 2002, pp. 206-223.

11. Al-Haj, A., A dual transform audio watermarking algorithm. Multimedia Tools and Applications, 73(3), 2014, pp. 1897-1912.

12. Peng, H. and J. Wang, Optimal audio watermarking scheme using genetic optimization. Annals of telecommunications-annales des télécommunications, 66(5-6), 2011, pp. 307-318.

13. Al-Haj, A., An imperceptible and robust audio watermarking algorithm. EURASIP Journal on Audio, Speech, and Music Processing, 2014(1), 2014, pp. 1-12.

14. Li, J., et al., A multipurpose audio aggregation watermarking based on multistage vector quantization. Multimedia Tools and Applications, 68(3), 2014, pp. 571-593.

15. Wang, Y., S. Wu, and J. Huang, Audio watermarking scheme robust against desynchronization based on the dyadic wavelet transform. EURASIP Journal on Advances in Signal Processing, 2010, p. 13.

16. Yan, D. and R. Wang, Huffman table swapping-based steganograpy for MP3 audio. Multimedia Tools and Applications, 52(2-3), 2011, pp. 291-305.

17. Xuemei, J., L. Quan, and W. Qiaoyan, A new video watermarking algorithm based on shot segmentation and block classification. Multimedia Tools and Applications, 62(3), 2013, pp. 545-560.

18. Li, Z., X.-W. Chen, and J. Ma, Adaptively imperceptible video watermarking based on the local motion entropy. Multimedia Tools and Applications, 2013, pp. 1-22.

19. Xu, D., R. Wang, and J. Wang, Prediction mode modulated data-hiding algorithm for H. 264/AVC. Journal of Real-Time Image Processing, 7(4), 2012, pp. 205-214.

20. Agarwal, H., R. Ahuja, and S. Bedi, Highly Robust and Imperceptible Luminance Based Hybrid Digital Video Watermarking Scheme for Ownership Protection. International Journal of Image, Graphics and Signal Processing (IJIGSP), 4(11), 2012, pp. 47.

21. Rathod Jigisha D, R.V.M., A HYBRID DWT-SVD METHOD FOR DIGITAL VIDEO WATERMARKING. International Journal of Advanced Research in Computer and Communication Engineering, 2(7), 2013.

22. Ejaz, N., et al., Adaptive image data hiding using transformation and error replacement. Multimedia Tools and Applications, 73(2), 2014, pp. 825-840.

23. Tsougenis, E., G. Papakostas, and D. Koulouriotis, Image watermarking via separable moments. Multimedia Tools and Applications, 2014, pp. 1-28.

24. Xiao, D., S. Hu, and H. Zheng, A high capacity combined reversible watermarking scheme for 2-D CAD engineering graphics. Multimedia Tools and Applications, 2013, pp. 1-18.

25. Cedillo-Hernandez, M., et al., Robust watermarking method in DFT domain for effective management of medical imaging. Signal, Image and Video Processing, 2013, pp. 1-16.

26. Chu, S.-C., et al., Genetic watermarking for zerotree-based applications. Circuits, Systems & Signal Processing, 27(2), 2008, pp. 171-182.

27. Huang, H.-C., C.-M. Chu, and J.-S. Pan, The optimized copyright protection system with genetic watermarking. Soft computing, 13(4), 2009, pp. 333-343.

28. Hsieh, M.-S. and D.-C. Tseng, Perceptual digital watermarking for image authentication in electronic commerce. Electronic Commerce Research, 4(1-2), 2004, pp. 157-170.

29. Qi, P., Impact analysis of digital watermarking on perceptual quality using HVS models, 2005, Citeseer.

30. Agrawal, R. and J. Kiernan. Watermarking relational databases. in Proceedings of the 28th international conference on Very Large Data Bases. 2002, VLDB Endowment.

31. Ullah, H. and A. Khan. Imperceptibility Analysis of Watermarked Database. in International conference on cultural technology, 2017. Chiang Mai, Thailand.

32. Agrawal, R., P.J. Haas, and J. Kiernan, Watermarking relational data: framework, algorithms and analysis. *The VLDB journal*, 12(2), 2003, pp. 157-169.

33. Johnson, R.A. and D.W. Wichern, *Applied multivariate statistical analysis*, 2014: Pearson.

34. Hair, J.F., et al., *Multivariate Data Analysis*, Pearson Prentice Hall. Upper Saddle River, NJ, 2006.

35. Ge, W. and G. Whitmore, Binary response and logistic regression in recent accounting research publications: a methodological note. *Review of Quantitative Finance and Accounting*, 2010. 34(1): p. 81-93.

36. Bewick, V., L. Cheek, and J. Ball, Statistics review 14: Logistic regression. *Critical Care*, 9(1), 2005, pp. 1-7.

Information about the authors:

Hameed Ullah - MSCS Student at Iqra University Islamabad Campus. This project implemented for final year thesis. Research interests are information security, cryptography, watermarking, encryption and computer network security.

Dr. Aihab Khan – working as Associate Professor at Iqra University Islamabad Campus. The project implemented under his supervision. His areas of interest are information security, watermarking, computer security, algorithms and software quality assurance.

Dr. Basheer Ahmad – working as Professor of Statistics and HOD of Management Sciences Department at Iqra University Islamabad Campus. His research interests are statistical analysis, statistical modelling, regression analysis and data analysis.

Manuscript received on 14 October 2017