
A NEW APPROACH FOR SPEECH AUDIO STEGANALYSIS USING DELAY VECTOR VARIANCE METHOD

*Emrah YÜRÜKLÜ**
*Osman Hilmi KOÇAL***
*Erdoğan DİLAVEROĞLU****

Abstract: We investigate the use of delay vector variance-based features for recorded speech steganalysis. Considering that data hiding within a speech signal distorts the properties of the original speech signal, we design a steganalyzer that uses surrogate data based delay vector variance (*DVV*) features to detect the existence of a stego-signal. We evaluate the performance of the proposed *DVV* features as steganalyzer with numerical results.

Keywords: Steganography, steganalysis, speech, chaos, false-neighbors, Lyapunov exponent, surrogate data, delay vector variance.

Konuşma Sesleri Steganalizine Gecikmeli Vektör Varyans Metodunu Kullanan Yeni bir Yaklaşım

Özet: Bu çalışmada gecikmeli vektör varyans metodunun, kayıt edilmiş konuşma seslerinin steganalizinde kullanımı araştırılmıştır. Konuşma sinyali içerisine veri gizlemenin orijinal ses verisinin özelliklerini bozacağı gerçeğinden göz önünde bulundurarak, vekil verileri kullanan gecikmeli vektör varyans (*DVV*) özelliklerindeki değişimi kontrol eden ve böylece stego-sinyalin varlığını tespit etmeye çalışan yeni bir steganalizör tasarlanmıştır. *DVV* özelliklerini baz alan önerilen steganalizörün performans başarımları ise nümerik değerlerle verilmiştir.

Anahtar Kelimeler: Steganografi, steganaliz, konuşma, kaos, hatalı-komşular, Lyapunov üsteli, vekil veri, gecikmeli vektör varyans.

1. INTRODUCTION

Steganography is the art and science of hiding the presence of communication by embedding secret messages into innocent-looking digital signals. To achieve secure and undetectable communication, *stego-signals*, containing a secret message, should be indistinguishable from *cover-signals*, signals not containing any secret message. In this respect, *steganalysis* is a set of techniques that aim to distinguish between cover-signals and stego-signals.

Audio signal steganalysis methods were proposed in the literature for various approaches. The work by Westfeld (2003) is for LSB based embedding, while the work by Fridrich and Goljan (2002) addresses the steganalysis of MP3stega algorithm. The steganalysis of LSB based

* Bursa Orhangazi University Engineering Faculty, Dept. of Electrical & Electronic Eng., Yıldırım 16310, Bursa.

** Yalova University, Engineering Faculty, Dept. of Computer Engineering, 77100 Yalova.

*** Uludağ University, Engineering Faculty, Dept. of Electrical & Electronic Engineering, Görükle 16059, Bursa.

İletişim Yazarı: O.H. Koçal (Osman.kocal@yalova.edu.tr)

embedding and Hide4pgp algorithm is explored by Johnson et al. (2005) spread spectrum watermarking and stochastic modulation steganography is considered by Westfeld and Pfitzmann (1999). Both watermarking and steganographic hiding methods in time and frequency domains are considered by Westfeld (2003). Chaotic type features are investigated by Koçal et al. (2008) and found very useful for distinguishing stego signals from cover signals.

Chaotic phenomena in speech is still being in subject of various researches. Though linearity and stationarity of the audio are inherent assumptions in many of the reported approaches, there is also theoretical and experimental evidence for the existence of chaotic phenomena in speech signals which linear modeling cannot cover it (Altun et al., 2005; Özer et al., 2003; Kokkinos and Maragos, 2005; Banbrook and McLaughlin, 1999). Delay vector variance (DVV) method which is based on surrogate time series (Theiler et al., 1999) is one of the techniques which can determine the level of nonlinearity in signals (Gautama et al., 2003). Assuming that data hiding leads to additive distortion on speech signal, it is anticipated that this process will change the nonlinear and chaotic structure of the speech signal, therefore DVV and chaotic-based features.

The change in chaotic structure means also the change in nonlinear structure, thus manifests as deviation from the numerical values of the DVV's of the cover signal, and therefore enables the detection of the possible existence of a hidden signal. The details of the data hiding effect on DVV will be given in the next section. By the motivation of the effect of data hiding on chaotic properties of speech signals, we propose a steganalysis technique that utilizes DVV features and the other chaotic features as well.

The rest of the paper is organized as follows: In Section 2 we give an overview of *DVV* features for speech signals and show how data hiding changes these features. DVV feature extraction, feature selection and experimental results are given in Section 3. Conclusions are drawn in Section 4.

2. DVV METHOD

For assessing the nonlinearity presence in time series, “surrogate time series” method is a widely-used technique which was offered by Theiler et al. (1992). Surrogate time series is produced with the same magnitude and the similar phase of Fourier transform of the original time series. There are some different approaches for producing surrogate time series, but the most widely used one is the method named iterative amplitude adjusted Fourier transform (iAAFT) which was proposed by Schreiber and Schmitz (1997). In this paper this approach is used for all produced surrogate time series. For detailed information please refer to Schreiber and Schmitz (1997).

By comparing original and surrogate time series' characteristics, it can be anticipated the level of nonlinearity in the time series. Some metrics based on surrogate times are defined for doing this comparison.

2.1. Nonlinearity Measurements

Besides the delay-vector-variance method, there are two other approaches to predict the nonlinearity of the time series in the study by Schreiber and Schmitz (1997); the third-order autocovariance and asymmetry due to time reversal (Gautama et al., 2003).

The third-order autocovariance is a higher-order extension of the traditional autocovariance and can be given by;

$$t^{C3}(\tau) = \langle x_n x_{n-\tau} x_{n-2\tau} \rangle \quad (1)$$

Here τ is the time lag. Given time series is accepted as time reversible, if the probabilistic properties don't change regarding to time reversal. Invariance for the probabilistic features can be measured by using the equation of

$$t^{REV}(\tau) = \left\langle (x_n - x_{n-\tau})^3 \right\rangle \quad (2)$$

In the studies by Gautama et al. (2003) and Schreiber and Schmitz (1997) has been shown that with combination of *DVV*, these two methods are very helpful for two-tailed tests for nonlinearity.

2.2. DVV-Method

This method was first proposed in 2003 (Gautama et al., 2003) and is a generic nonlinearity predictor which uses phase spaces of time series. For a given embedding dimension, D_E , set of delay vectors, $\mathbf{x}(n)$, are generated with time lag, τ ,

$$\mathbf{x}(n) = [x_{n-D_E\tau}, \dots, x_{n-\tau}] \quad (3)$$

Defining the proper values and the effects of selecting improper values for the embedding dimension, D_E , and time delay, τ , is widely described in [6] and [10]. For a given embedding dimension, D_E , DVV method computes the mean target variance, σ^{*2} , for all of the sets of Ω_n . Here every Ω_n is group which is consisted of those delay vectors which are in a certain distance to $\mathbf{x}(n)$. The certain distance is varied regarding to the distribution of pairwise distances between delay vectors (Gautama et al., 2003).

DVV method can be summarized as follow;

- For the given embedding dimension, D_E ; the mean, μ_d , and the standard deviation, σ_d , are computer aver the all pairwise Euclidean distances between delay vectors,
- For the given embedding dimension, D_E ; $\Omega_n(r_d)$ sets are generated according to this formula $\Omega_n(r_d) = \{\mathbf{x}(i) \mid \|\mathbf{x}(n) - \mathbf{x}(i)\| \leq r_d\}$; $i=1,2,\dots,N$. Here N is the sample count in time series of \mathbf{x} and r_d is the certain distance which has to be chosen from the interval $[\mu_d - n_d\sigma_d; \mu_d + n_d\sigma_d]$, where n_d is a parameter that controls the span over which to compute the *DVV*-plot.
- For the given embedding dimension, D_E ; the variance of every set of $\Omega_n(r_d)$, $\sigma_n^2(r_d)$, is computed. The average value of all sets $\Omega_n(r_d)$ is normalized by the variance of the time series, σ_x^2 . At the end the measure of the unpredictability is computed by;

$$\sigma^{*2}(r_d) = \frac{1}{N} \sum_{n=1}^N \sigma_n^2(r_d) \quad (4)$$

Unless one set, $\Omega_n(r_d)$, contains more than 30 delay vectors, it is not taken into account while doing the computations (Gautama et al., 2003).

As a result of doing standardization of the distance axis, the DVV-plots are really straightforward to interpret. In fig.1 DVV-plots of four benchmark signals are illustrated as a sample to

explain DVV better. According to the plots, Henon Map signal (A) is the most predictable signal while colored noise (C) is the most unpredictable signal. Note that, all the x-axis are standardized as distances, so it allows to make the comparison easier (Gautama et al., 2004).

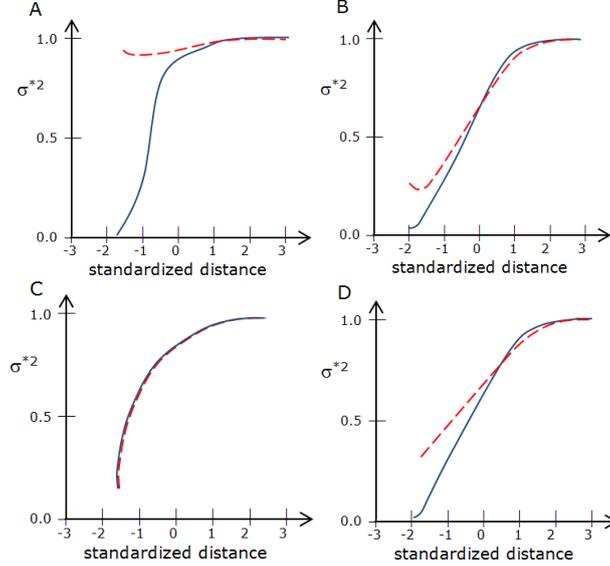


Figure 1:

Solid curves represent the DVV-plots for the Hénon Map (A), Mackey-Glass (B), colored noise (C), and the laser time series (D). Average DVV-plots computed over 99 surrogates are shown as dashed curves.

3. USING DVV AS A STEGANALYZER FEATURE

With evaluating fig.1, we can have ideas about the nonlinearity dynamics of the signal. But, this approach is not enough as long as we can't have objective decisions every time that we make evaluations. Because of this reason, Schreiber and Schmitz have constituted the below method which produces numerical results (Schreiber and Schmitz, 1997);

1. The number of surrogate data, n_{surr} , is decided according to defined limit of significance, α (generally 0.05),

$$n_{surr} \geq \frac{2}{\alpha} - 1 \quad (5)$$

2. n_{surr} surrogate time series are produced from original time series; $s_1, s_2, \dots, s_{n_{surr}}$,
3. DVV curves of original and surrogate time-series are constructed,
4. The difference between DVV curves of original and surrogate are calculated as $RMSE$ ($RMSE_1, RMSE_2, \dots, RMSE_{n_{surr}}$.)
5. For every original time series (sound record), feature vector is constituted with mean, variance, skewness and kurtosis values calculated over $RMSE$ values of surrogate time series;

$$\mathbf{F}_{surr} = \{mean(RMSE) \ var(RMSE) \ ske(RMSE) \ kur(RMSE)\} \quad (6)$$

In Ω_n set, the square value of the neighborhood distances between all delayed vector couples is;

$$d(x(n), x(m))^2 = \|x(m) - x(n)\|^2 = \sum_{k=1}^{D_E} (x(n+k) - x(m+k))^2 \quad (7)$$

The mean, μ_d , and variance values, σ_d^2 , in Ω_n set are:

$$\begin{aligned} \mu_d &= E[d] \\ x(m) \in \Omega_n & \\ \sigma_d^2 &= E[(d - \mu_d)^2] = E[d^2] - \mu_d^2 \\ x(m) \in \Omega_n & \end{aligned} \quad (8)$$

Using the assumption of hiding data into cover-signals means that adding zero-mean and σ_ε^2 variances white noise with magnitude of $\varepsilon(n)$, if we reorganize the equation (7) for both cover and stego-signals;

$$d_c(x(n), x(m))^2 = d_c^2 = \sum_{k=1}^{D_E} (x(n+k) - x(m+k))^2 \quad (9)$$

$$d_s(x(n), x(m))^2 = d_s^2 = \sum_{k=1}^{D_E} ([x(n+k) + \varepsilon(n+k)] - [x(m+k) + \varepsilon(m+k)])^2 \quad (10)$$

If we expand the term of d_s^2 which is used for stego-signals;

$$\begin{aligned} d_s^2 &= \sum_{k=1}^{D_E} \{x^2(n+k) + 2x(n+k)\varepsilon(n+k) + \varepsilon^2(n+k) + x^2(m+k) + 2x(m+k)\varepsilon(m+k) + \varepsilon^2(m+k) \\ &\quad - 2[x(n+k)x(m+k) + x(n+k)\varepsilon(m+k) + \varepsilon(n+k)x(m+k) + \varepsilon(n+k)\varepsilon(m+k)]\} \end{aligned} \quad (11)$$

If we put $d_c^2 = \sum_{k=1}^{D_E} x^2(n+k) + 2x(n+k)x(m+k) + x^2(m+k)$ term in eq.(11);

$$\begin{aligned} d_s^2 &= d_c^2 + \sum_{k=1}^{D_E} \{2x(n+k)\varepsilon(n+k) + \varepsilon^2(n+k) + 2x(m+k)\varepsilon(m+k) + \varepsilon^2(m+k) \\ &\quad - 2[x(n+k)\varepsilon(m+k) + \varepsilon(n+k)x(m+k) + \varepsilon(n+k)\varepsilon(m+k)]\} \end{aligned} \quad (12)$$

As long as all the pairs of $\langle \varepsilon(m+k), \varepsilon(n+k) \rangle$, $\langle \varepsilon(m+k), x(m+k) \rangle$, $\langle \varepsilon(m+k), x(n+k) \rangle$, $\langle \varepsilon(n+k), x(m+k) \rangle$ ve $\langle \varepsilon(n+k), x(n+k) \rangle$ are independent identically distributed, the mean values of eq.(12) can be simplified as;

$$\begin{aligned} E[d_s^2] &= E[d_c^2] + D_E \cdot E[\varepsilon^2(n+k)] + D_E \cdot E[\varepsilon^2(m+k)] \\ x(n) \in \Omega_n \quad x(n) \in \Omega_n & \end{aligned} \quad (13)$$

Because of $\varepsilon(n)$ is zero-mean and σ_ε^2 varianced white noise;

$$\begin{aligned} E[d_S^2] &= E[d_C^2] + D_E \cdot 2\sigma_\varepsilon^2 \\ x(k) \in \Omega_n \quad x(k) \in \Omega_n \end{aligned} \quad (14)$$

The variance value which is used in DVV analysis and seen in eq.(8) can be defined for cover and stego-signals as below;

$$(\sigma_d^2)_C = E[d_C^2] - (\mu_d)_C^2 \quad (15)$$

$$(\sigma_d^2)_S = E[d_S^2] - (\mu_d)_S^2 \quad (16)$$

By using eq.(14) and (15), the term $(\sigma_d^2)_S$ can be written as;

$$(\sigma_d^2)_S = (\sigma_d^2)_C + (\mu_d)_C^2 + D_E \cdot 2\sigma_\varepsilon^2 - (\mu_d)_S^2 \quad (17)$$

With the assumption of $(\mu_d)_C^2 \approx (\mu_d)_S^2$ eq.(17) can be simplified as;

$$(\sigma_d^2)_S = (\sigma_d^2)_C + D_E \cdot 2\sigma_\varepsilon^2 \quad (18)$$

With this result, we can say that, variance values of stego-signals which are used in DVV analysis are always higher than the cover-signals'.

First three elements of the proposed F_{surr} feature vector can be seen in fig.2 which are calculated over 2000 stego and cover-signals with using DSSS (Bender et al., 1996) and Stochastic Modulation (Fridrich and Goljan, 2003) steganography techniques. Here n_{surr} value is 20 and the difference between feature values of cover and stego-signals can be easily seen from the figure.

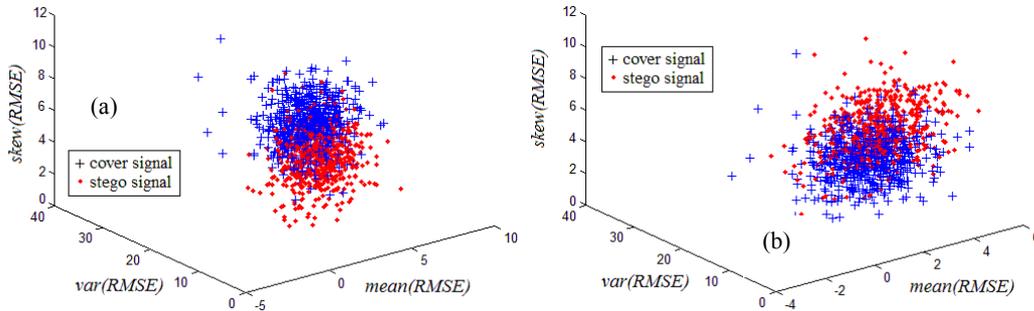


Figure2:

The first 3 elements of F_{surr} of 2000 cover and stego-signals which are constituted with steganographic techniques DSSS (a) and Stochastic Modulation (b). n_{surr} is 20 for F_{surr} .

4. EXPERIMENTAL RESULTS

We have performed tests over nine different data hiding methods. Here five of them are watermarking techniques, while the other four are steganographic techniques. The selection of watermarking methods is to extend the case to the widest possibilities, where a stego-signal can be deliberately changed before the receiver gets it (Simmons, 1984). The selected watermarking techniques are direct-sequence spread spectrum (DSSS), frequency hopping with spread spectrum (FHSS) and echo hiding (ECHO) by Bender et al. (1996) and DCT based watermarking method (COX) by Cox et al. (1997). Four steganographic methods are Steganos (2011), MP3Stego (2011), Steghide (2011), Hide4Pgp (2011) and Stochastic Modulation (Fridrich and Goljan, 2003). The rationality of using them is their popularity, free availability, and wide usage in related works. The performances for individual embedding methods are determined for to see whether the proposed feature set is useful for audio steganalysis.

4.1. Data Set

The database used for the test scenario is a subset of TIMIT speech database (2011), which is widely known in the evaluation of automatic speech recognition systems. TIMIT database is constituted of over 6300 utterances of 630 male-female speakers sampled at 16 KHz. 2000 speech segments are randomly selected from TIMIT database for our experiments, ignoring dialects and male-female differences. For every single data hiding method, stego-signal subset is constituted with embedding data into utterances with using the data hiding method.

4.2. Embedding

We have embedded messages into these excerpts with nine embedding methods. The procedure of embedding messages is randomly selecting half of the set of the stego and cover-signals for training, and leaving the other 50% for testing. An objective distortion measure is defined, Signal to Watermark Ratio (SWR), to define the payload level in watermarking methods:

$$SWR = \frac{\sum x(n)^2}{\sum (x(n) - y(n))^2} \quad (19)$$

where $x(n)$ and $y(n)$ signify cover and watermarked signals, respectively. For steganographic methods, for three different embedding rates performance of each individual method are evaluated: 100% and 50% of maximum allowed capacity.

4.3. Feature Set

The feature set proposed in the study has 4 elements which are all based on DVV-values as described in eq.(6). To see the whole performance of chaotic based features, also a whole feature set based on chaotic features, which consists 26 elements, is defined as below;

$$\mathbf{F}(D_E) = \{\mathbf{F}_{FNF}(D_E) | D_E = 3,4,5,6,7\} \cup \{\lambda_i | i = 1,2,3,4,5,6,7\} \cup \mathbf{F}_{Surr} \quad (20)$$

Here $\mathbf{F}_{FNF}(D_E)$ is the feature set of FNF (False Neighbour Fraction) which is based on the statistical values of FNF method (Kennel and Abberbanel, 2002), while λ_i is i . LE value (Lyapunov Exponent) of the signal (Hilborn, 2000). Using these chaotic features in audio steganalysis is very advantageous as described by Koçal et al. (2008). $\mathbf{F}_{FNF}(D_E)$ is described as;

$$\mathbf{F}_{FNF}(D_E) = [FNF, \text{mean}(d_{D_E}(\mathbf{s}(n), \mathbf{s}(m))), RMS(d_{D_E}(\mathbf{s}(n), \mathbf{s}(m)))] \quad (21)$$

where three elements of feature vector are the fraction of false neighbors, the average size of the neighborhood, and the root-mean-squared size of the neighborhood. The FNF and the LE values of the signals are calculated with TISEAN (Hegger et al., 1999) software package.

4.4. Feature Selection and Classifiers

For achieving highest detection rate to evaluate redundant features, sequential forward floating search method (SFFS) by Pudil and Novovicova, (1994) is coupled with support vector machine (SVM) classification method (OSU-SVM Matlab Toolbox, 2011). Radial basis function with Gamma = 4 is used in SVM as SVM function.

4.5. Simulation Results

For assessing the validity of proposed feature set, several tests have been carried out also considering the effect of payload. The performance of DVV based steganalyzer is given in Table I for TIMIT dataset. As performance metrics; results are given in the percentage of the performance, Missed Detection count (MISS) and False Alarm count (FA) as there are 1000 samples in performance tests.

By Özer et al. (2006), it is observed that SWR figures greater than 38dB for DSSS, 34dB for FHSS, 20dB for Cox, and 18dB for ECHO, become noticeable, namely audible. For this reason 20dB, 30dB and 40dB are selected as payloads. In detail, it can be seen that 30 dB payload causes performance drops to 55% in Cox method; but note that this level of payload is audible. For steganographic methods performance of each individual method are evaluated for two different embedding rates: 100% and 50% of maximum allowed capacity. There is no audibility concern for steganographic methods, because the maximum embedding rate used in the tests is equal to maximum allowed capacity which is the audibility threshold. As expected, performance decreases as payload decreases.

Table 1. SVM-based steganalysis performance for different embedding capacities with DVV features

WATERMARKING METHODS					STEGANOGRAPHY METHODS				
Method	SWR	MISS	FA	%	Method	Capacity Usage	MISS	FA	%
DSSS	20dB	0/1000	0/1000	100	STEGA	%50	142/1000	166/1000	84,5
	30 dB	1/1000	2/1000	99,9		%100	64/1000	73/1000	93,2
	40 dB	5/1000	6/1000	99,5	HIDE4PGP	%50	42/1000	176/1000	89,1
FHSS	20 dB	2/1000	0/1000	99,9		%100	33/1000	84/1000	94,2
	30 dB	5/1000	6/1000	99,5	STEGHIDE	%50	121/1000	154/1000	86,3
	40 dB	16/1000	13/1000	98,0		%100	48/1000	83/1000	93,5
ECHO	20 dB	98/1000	236/1000	83,3	MP3	%50	287/1000	167/1000	77,3
	30 dB	176/1000	395/1000	71,5		%100	203/1000	124/1000	83,7
	40 dB	293/100	489/1000	60,9					
COX	20 dB	285/1000	212/1000	75,2					
	30 dB	452/1000	457/1000	54,6					
	40 dB	478/1000	485/1000	51,9					
STOMOD	20 dB	13/1000	8/1000	99,0					
	30 dB	53/1000	34/1000	95,7					
	40 dB	171/1000	98/1000	89,8					

5. CONCLUSION

As conclusion, steganalyzers based on the proposed DVV features has high performances regarding mostly known steganography methods (Johnson et al., 2005; Koçal et al., 2008; Özer et al., 2003). If DVV-based features are used with the other chaotic based features, performance of the steganalyzer becomes more successful according the DVV-based steganalyzer. With these results it can be easily said that DVV-based features can be used as speech audio steganalysis and also for nonlinearity detection as well.

REFERENCES

1. Westfeld, A.P. (2003). Detecting Low Embedding Rates, *Fabien A. P. Petitcolas (Ed.): Information Hiding. 5th International Workshop, IH 2002 Noordwijkerhout*, pp. 324–339, Springer-Verlag Berlin Heidelberg..
2. Fridrich, J. and Goljan, M. (2002). Practical Steganalysis of Digital Images - State of the Art, *Proc. SPIE Photonics West*, Vol. 4675, pp. 1—13.
3. Johnson, M.K., Lyu, S. and Farid, H. (2005). Steganalysis of Recorded Speech, *Proc. SPIE Symposium on Electronic Imaging*, San Jose, CA.
4. Westfeld, A.P. and Pfitzmann, A. (1999). Attacks on steganographic systems, *Information Hiding, LNCS 1768*, pp. 61-66, Springer-Verlag Heidelberg.
5. Koçal, O.H., Yürüklü, E. and Avcıbaşı, İ. (2008). Chaotic-type features for Speech Steganalysis, *IEEE Transactions on Information Forensics and Security*, Vol. 3, Issue 4, pp. 651-661.
6. Altun, O., Sharma, G., Celik, M., Sterling, M., Titlebaum, E. and Bocko, M. (2005). Morphological Steganalysis of Audio Signals and the Principle of Diminishing Marginal Distortions”, *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Proceedings. (ICASSP '05)*, Vol. 2, pp. 21- 24.
7. Ozer, H., Avcibas, I., Sankur, B. and Memon, N. (2003). Steganalysis of Audio Based on Audio Quality Metrics, *Proc. SPIE, Security and Watermarking of Multimedia Contents V*, Vol. 5020, pp. 55-66.
8. Kokkinos, I. and Maragos, P. (2005). Nonlinear Speech Analysis Using Models for Chaotic Systems, *IEEE Transactions on Speech and Audio Processing*, Vol. 13, No. 6, pp. 1098-1109.
9. Banbrook, M. and McLaughlin, S., (1999). Speech Characterization and Synthesis by Nonlinear Methods”, *IEEE Transactions on Speech and Audio Processing*, Vol. 7, No. 1, pp. 1-17.
10. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J. D. (1992). Testing for Nonlinearity in Time Series: The Method of Surrogate Data, *Physica D* 58. pp. 77 - 94.
11. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2003). Indications of nonlinear structures in brain electrical activity, *Physical Review E*, Vol. 67, No. 046204.
12. Schreiber, T. and Schmitz, A. (1997). Discrimination power of measures for nonlinearity in a time series, *Phys. Rev. E, Vol:55*. pp. 5443-5447.
13. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2004). A Novel Method for Determining the Nature of Time Series, *IEEE Transactions On Biomedical Engineering*, Vol. 51, No. 5.
14. Simmons, G. J. (1984). Prisoners' Problem and the Subliminal Channel (The), *CRYPTO83 - Advances in Cryptology*, pp. 51-67.

15. Bender, W., Gruhl, D., Morimoto, N. and Lu, A. (1996). Techniques for data hiding, *IBM Systems Journal*, vol. 35, no: 3&4, pp. 313-336.
16. Fridrich, J. and Goljan, M. (2003). Digital Image Steganography Using Stochastic Modulation, *Proc. SPIE Electronic Imaging*, Santa Clara, CA, pp. 191-202.
17. Cox., I. J., Kilian, J., Leighton, F.T. and Shamoon, T. (1997). Secure spread spectrum watermarking for multimedia, *IEEE Tr. Image Proc.*, vol 6, pp. 1673 – 1687.
18. www.steganos.com, Access Date: 23.Nov.2011, Subject: Steganos.
19. www.petitcolas.net/fabien/ steganography/mp3stego, Access Date: 12.Sep.2011, Subject: Mp3stego.
20. http://steghide.sourceforge.net, Access Date: 13.Sep.2011, Subject: S. Hetzl.
21. www.heinz-repp.onlinehome.de/Hide4PGP.htm, Access Date: 12.Sep.2011, Subject: Hide4PGP
22. <http://www.mpi.nl/world/tg/corpora/timit/timit.html>, Access Date: 05.Jul.2011, Subject: TIMIT Speech Database
23. Kennel, M. B. and Abarbanel, H. D. I. (2002). False neighbors and false strands: A reliable minimum embedding dimension algorithm”, *Physical Review E*, Vol. 66, 026209.
24. Hilborn, R. (2000). *Chaos and Nonlinear Dynamics*, Oxford University Press, 2. ed.
25. Hegger, R., Kantz, H. and Schreiber, T. (1999). Practical Implementation of Nonlinear Time Series Methods: The TISEAN Package, *Chaos*, Vol. 9, pp. 413-435.
26. Pudil, P., Novovicova, J. and Kittler, J. (1994). Floating Search Methods in Feature Selection, *Pattern Recognition Lett.* Vol 15, pp. 1119-1125.
27. http://eewww.eng.ohio-state.edu/~maj/osu_svm/, Access Date: 24.Sep.2011, Subject: OSU-SVM Matlab Toolbox.
28. Özer, H., Sankur, B., Memon, N. and Avcıbaşı, İ. (2006). Detection of Audio Covert Channels Using Statistical Footprints of Hidden Messages, *Digital Signal Processing*, Vol.16, pp. 389-401.

Makale 19.11.2013 tarihinde alınmış, 07.12.2013 tarihinde düzeltilmiş, 09.12.2013 tarihinde kabul edilmiştir.