SPEECH ENCRYPTION USING CIRCULANT TRANSFORMATIONS

G. Manjunath and G.V. Anand

Department of Electrical Communication Engineering,
Indian Institute of Science, Bangalore, 560012, India.
E-mail ***addresses :*** manju@protocol.ece.iisc.ernet.in, anandgv@ece.iisc.ernet.in

ABSTRACT

A new analog encryption technique using unitary circulant transformations of the sampled analog signal **is** proposed. **This** technique is less vulnerable to attack than other **known** schemes based on transposition(scrambling) of the signal components in time or frequency domain. The new technique introduces primarily a phase distortion. A new distonion measure capable of quantifying phase distortion is proposed. ***An*** optimal speaker-specific key generation scheme is then developed for maximizing an objective function based on the new distortion measure.

Key words: Analog speech encryption, key generation, distortion measure.

1. INTRODUCTION

Analog encryption schemes [1] are relatively insecure compared to digital encryption schemes, but have an advantage in that they can be easily inter-faced with the existing channels such as telephone, satellite and mobile communication links. The main reasons for the vulnerability of analog encryption schemes are the bandwidth constraints of the analog channels, the nature of the speech signal and the various encryption schemes which involve permutations as a key to hide the signal. Compared to the time-domain scrambling algorithms, the transform domain speech scramblers have a distinct advantage in that the number of effective permutations is much larger than the number of permutations available in the time-domain scrambling algorithms. Hence, the security offered by transform domain scramblers against a brute force attack which tries out the permutations is higher and hence they are prevalent. Such an attack on the transform domain scramblers requires the cryptanalyser to listen to each of the signals obtained after repositioning the transform coefficients followed by the inverse transform, which is wearisome. Goldburg et al. [2] in their paper propose an attack on the transform domain scramblers without a human intervention at each stage. **Using** spectral matching techniques, an estimate of the permutation is determined in [2]. This suggests that an attack can be done with the aid of computational facilities rather than by listening and determining the content of speech for each guess of the key used. Thus transpositions alone are not enough to prevent an attack of this kind. In the next Section we propose an encryption scheme that does not use transpositioning of speech data.

2. ENCRYITION USING CIRCULANTS

Let a finite frame of speech samples of length N be represented hy the vector $\mathbf{x} = [x(0)\, x(1)\ \ldots x(N-1)]^T$. Consider the set

$$\Omega = \left\{\mathbf{h}\ :\ |H(k)| = 1,\ k = 0,1,\ldots,N-1;\ \mathbf{h} \in \Re^N\right\},\quad (1)$$

where $H(6)$ denotes the DFT coefficients of the vector **h.** Let $\phi(k)$ denote the angle or phase of $H(k)$. For **h** to be an element of the set Ω, each of the principal phase values $\phi(1),\ \phi(2),\ \ldots,\ \phi\left(\frac{N}{2}-1\right)$ may be chosen arbitrarily and independently from the interval $[-\pi,\pi]$, while $\phi(0)$ and $ ($) may be equal to either 0 or π. The other phase values are fixed by the conjugate symmetry property of the DFT of a real sequence. The cardinality of the set ll is infinite and it **is** uncountable. We define the vector of encrypted speech samples as $\tilde{\mathbf{x}} := \mathbf{h} \otimes \mathbf{x}$, where $\otimes$ denotes circular convolution. We can also write the same relation as $\tilde{\mathbf{x}} = \mathbf{a} \odot \mathbf{x}$, where $\odot$ denotes circular correlation and the elements of the vector **a** are given by $a(n) := h((-n) \bmod \mathrm{N})$, $\mathbf{n} = 0, 1, \ldots, \mathbf{N} - 1$. Equivalently in matrix notation $\tilde{\mathbf{x}} = \mathbf{Ax}$, where $\mathbf{A}$ is a circulant matrix [3] whose rows are the vector **a** and its circular shifts. From the definition of $a(n)$, the DFT relation $a(n) \odot a(n) \rightleftharpoons H(k) \cdot H^*(k)$ follows. Since $H(k)\ .H^*(k) = [1\ 1\ \ldots,\ 1]^T$, and its inverse DFT is the vector $[1\ 0\ 0,\ \ldots,\ 0]^T$. Writing this is an equation, we have

$$\sum_{m=0}^{N-1} a(m)a((m-\mathbf{n}) \bmod \mathrm{N}) = \delta(n). \quad (2)$$

Equation **(2)** indicates that the vector **a** and its $\mathbf{N}-\mathbf{1}$ circular shifts form an orthonormal basis of $\Re^N$. For each $\mathbf{h} \in$ 11, we get a unique circulant $\mathbf{A}$ and hence there are infinitely many NxN onhonormal circulants which can be used as the keys for encryption of the speech vector **x.** Since $\mathbf{A}$ is orthonormal. the original signal (i.e. decrypted signal) can be recovered by the transformation $\mathbf{x} = \mathbf{A}^T\tilde{\mathbf{x}}$.

3. ANALYSIS OF THE ENCRIYITION SCHEME

The circulant transformation produces a redistribution of energy in the temporal domain rather than only a transposition of the speech data as in the conventional time-domain scrambling schemes. Since the redistribution of energy is throughout and within the expanse of the frame, greater lengths of the frame can engender higher uinitelligibility than those engendered by smaller lengths. The relative importance of phase in a **signal** as compared to that of magnitude increases as the length of the signal under consideration increases **[4].** Hence when the signal length is large enough, a random phase distortion can result in serious impairing of the temporal structure of the speech signal. The original and the encrypted signal for the utterance **"*while** attention*" by an adult male speaker is shown in Figure 1. A frame length of $\mathbf{N} = \mathbf{512}$ was used for the signal sampled at 8 kHz.

The perceptual distonion caused hy the process of encryption can he explained as the result of destroying the formant transitions which are important acoustic cues for the perception of speech. Libermann et al. [5] show the necessity of these acoustic cues for the perception of phonemes. Speech is a complex and varied activity encompassing not just words, gestures and tone of voice

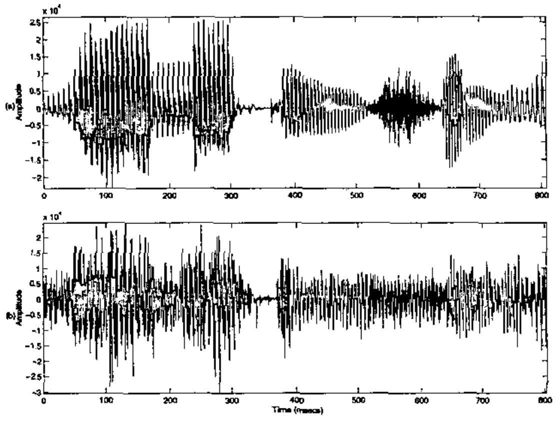

Fig. 1. (a) The original discrete signal for the utterance *'while attention'* in time domain. (b) the corresponding scrambled signal in time domain.

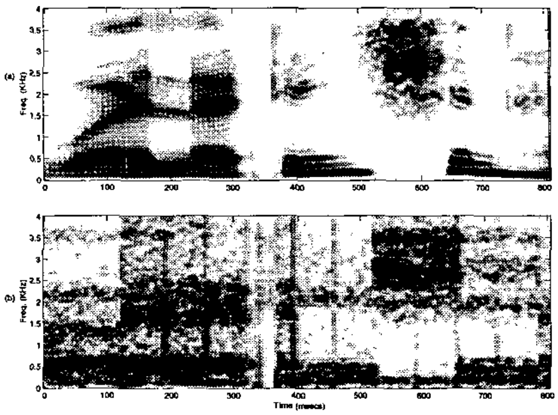

Fig. 2. (a) Spectrogram of the original signal. (b) Spectrogram of the encrypted signal.

but also silence; the near silence during the glottal closure period **during** which the speech signal has law amplitude values is **also** affected by the temporal distortion brought about by the encryption process. The silence between words in the original speech signal also contributes to the unintelligibility of the encypted signal. Figure 2 (a) and (b) show the spectrograms of the original and the encrypted signal for the same utterance *'while attention'* as that in Figure 1. It is evident from the spectrograms in Fig. ure 2 that the temporal redistribution of energy due to encryption has rendered the original formants distorted and introduced many new ones.

We now consider the analysis in the spectral domain. We had defined $\tilde{\mathbf{x}} = \mathbf{h}\otimes\mathbf{x}$. This is equivalent to multiplication in the DFT domain and therefore the k^{th} DFT coefficient of $\tilde{\mathbf{x}}$ is given by $X(k)e^{j\phi(k)}$. The Discrete Time Fourier Transforms of (DTFT) $\tilde{X}_c(\omega)$ of the signal $\tilde{\mathbf{x}}$ can be written as

$$\tilde{X}_c(\omega) = \sum_{k=0}^{N-1} X(k)e^{j\phi(k)}Y(\omega - \frac{2\pi}{N}k), \tag{3}$$

where the function $Y(\omega)$ is the interpolation function with the property that $Y(\frac{2\pi}{N}k) = \delta(k)$ [6]. If $\phi(k) = 0 \forall k$ in (3), we get an expression for $X_c(\omega)$, the DTFT of x. Hence from (3) we can observe that $\left|\tilde{X}_c(\omega)\right| \neq |X_c(\omega|$ except at the finite number of points at w $= \frac{2\pi}{N}i$, $i = 0, 1, ..., N - 1$. On the other-hand, the phase spectra of $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are different even at these points. Thus the unitary circulant transformation distorts the magnitude spectrum of the signal everywhere except at the N points mentioned above, and also the entire phase spectrum. It is obvious that as N increases, the distortion of the magnitude spectrum decreases. For typical frame lengths N such as 256 or more, one can neglect the bandwidth expansion due to encryption. It can be shown that bandwidth expansion due to encryption and the consequent distortion during transmission over a bandlimited channel is negligible [7].

4. KEY GENERATION SCHEME

An infinite choice of keys is available for encryption using circulants, each key corresponding to a phase vector $\Phi = \left[\phi(1)\,\phi(2)\,\ldots\,\phi\left(\frac{N}{2}-1\right)\right]^T$ subject to the constraint $\phi(0) = 0$ or π, $\phi\left(\frac{N}{2}\right) = 0$ or π, and $\phi\left((N-n) \bmod N\right) = \phi(n)$. We shall now address the issue of key selection that meets the twin requirements of high distortion of the encrypted signal and low vulnerability to attack. Most measures of speech quality described in the literature [8] primarily quantify spectral magnitude distortion. These measures are not suitable for quantifying the phase distortion effected by our encryption procedure. A widely used distortion measure is the ltakura distortion measure [8, 9] defined as

$$\ln d_I(\mathbf{x},\tilde{\mathbf{x}}) = \ln\left(\frac{\alpha^T\tilde{\mathbf{R}}\alpha}{\tilde{\alpha}^T\tilde{\mathbf{R}}\tilde{\alpha}}\right) = \ln\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{|X_c(\omega)|^2}{\left|\tilde{X}_c(\omega)\right|^2}d\omega\right], \tag{4}$$

where $\tilde{\mathrm{R}} = \mathrm{E}\left[\tilde{\mathrm{x}}\tilde{\mathrm{x}}^T\right]$ is the correlation matrix of $\tilde{\mathrm{x}}$, and $\boldsymbol{a} = [1\,\alpha(1)\,\alpha(2)\,\ldots\,\alpha(N-1)]^T$ and $\tilde{\boldsymbol{a}} = [1\,\tilde{\alpha}(1)\,\tilde{\alpha}(2)\,\ldots\tilde{\alpha}(N-1)]^T$ are respectively the vectors of linear predictor coefficients (LPCs) of x and $\tilde{\mathrm{x}}$. It is evident from Eq. (4) that the ltakura distance measure fails to quantify phase distortion. Hence we propose the following modified ltakura (MI) distortion measure :

$$\ln d_{MI}(\mathbf{x},\tilde{\mathbf{x}}) = \ln\left(\frac{\alpha_w^T\tilde{\mathbf{R}}_w\alpha_w}{\tilde{\alpha}_w^T\tilde{\mathbf{R}}_w\tilde{\alpha}_w}\right) + \ln\left(\frac{\alpha_z^T\tilde{\mathbf{R}}_z\alpha_z}{\tilde{\alpha}_z^T\tilde{\mathbf{R}}_z\tilde{\alpha}_z}\right), \tag{5}$$

where $\tilde{\mathrm{R}}_w = \mathrm{E}\left[\tilde{w}\tilde{w}^T\right]$, $\tilde{\mathrm{R}}_z = \mathrm{E}\left[\tilde{z}\tilde{z}^T\right]$, $\tilde{w}(n) := \frac{1}{2}\left(\tilde{x}(n) + \tilde{x}\left((N-n) \bmod N\right)\right)$, $\tilde{z}(n) := \frac{1}{2}\left(z(n) - z\left((N-n) \bmod N\right)\right)$, $w(n)$ and $z(n)$ are defined similarly, and α_w, $\tilde{\boldsymbol{a}}_w$, α_z and $\tilde{\alpha}_z$ are the **LPC** vectors of $\mathbf{w}$, $\tilde{\mathbf{w}}$, $\mathbf{z}$ and $\tilde{\mathbf{z}}$ respectively. We note that, at $\{w = \frac{2\pi}{N}k,\ k = 0, 1, ..., N-1\}$, DTFT of $\mathbf{w}$ $= \mathrm{Re}\left[X_c(\omega)\right]$ and DTFT of $\mathbf{z} = j\mathrm{Im}\left[X_c(\omega)\right]$ where Re[ ] and Im[ ] denote, respectively, the real and imaginary parts of their arguments. Hence for large N , the MI distortion measure is well approximated by

$$\ln d_{MI}(\mathbf{x},\tilde{\mathbf{x}}) \simeq \ln\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\mathrm{Re}\left[X_c(\omega)\right]}{\mathrm{Re}\left[\tilde{X}_c(\omega)\right]}\right)^2 d\omega\right]$$

$$+ \quad \ln\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\mathrm{Re}\left[X_c(\omega)\right]}{\mathrm{Im}\left[\widetilde{X}_c(\omega)\right]}\right)^2 d\omega\right]. \qquad (6)$$

Thus the MI distonion measure incorporates the effects of both magnitude distortion and phase distortion. Using linear prediction theory, it can be readily shown [7, 10] that $\left(\widetilde{\alpha}_w^T\widetilde{\mathbf{R}}_w\widetilde{\alpha_w}\right)^{-1} = \mathbf{u}^T\widetilde{\mathbf{R}}_w^{-1}\mathbf{u}$ and $\left(\widetilde{\alpha}_z^T\widetilde{\mathbf{R}}_z\widetilde{\alpha_z}\right)^{-1} = \mathbf{u}^T\widetilde{\mathbf{R}}_z^{-1}\mathbf{u}$, where $\mathbf{u} = [1\ 0\ 0\ \ldots 0]^T$is a Nx1 vector. Hence Eq. *(5)* can be written as

$$d_{MI}(\mathbf{x},\widetilde{\mathbf{x}}) = \alpha_w^T\widetilde{\mathbf{R}}_w\alpha_w\ \mathbf{u}^T\widetilde{\mathbf{R}}_w^{-1}\mathbf{u}\ \alpha_z^T\widetilde{\mathbf{R}}_z\alpha_z\ \mathbf{u}^T\widetilde{\mathbf{R}}_z^{-1}\mathbf{u}. \qquad (7)$$

For encryption using circulants, we have $\widetilde{\mathbf{x}} = \mathbf{A}\mathbf{X}$, and it can be shown that $\widetilde{\mathrm{R}}_w = \frac{1}{4}\left(\widetilde{\mathbf{R}} + \widetilde{\mathbf{R}}\mathbf{S} + \mathrm{S}\widetilde{\mathrm{R}} + \mathbf{S}\widetilde{\mathbf{R}}\mathbf{S}\right)$ and 6, $= \frac{1}{4}\left(\widetilde{\mathbf{R}} - \widetilde{\mathbf{R}}\mathbf{S} - \mathrm{S}\widetilde{\mathrm{R}} + \mathrm{S}\widetilde{\mathrm{R}}\mathrm{S}\right)$, where $\widetilde{\mathrm{R}} = \mathrm{ARA}^T$ and $\mathrm{R} = \mathrm{E}\,[\mathbf{xx'}]$. and S is a NxN symmetric permutation matrix. The matrix A = A **(Q)** depends **on** the phase vector Φ which is the key to encryption. The function in Eq. (7) is continuous and its domain is a compact set [12] and hence the maximum of the function is attained. The optimal key may **be** defined as the vector Φ that maximizes $\mathrm{d}(\Phi;\mathrm{R}) := d_{MI}(\mathbf{x},\widetilde{\mathbf{x}})$. But this is not a useful optimality criterion since the optimal key $\Phi_o = \Phi_o(\mathbf{R})$ is frame dependent. We therefore define a speaker-specific but frame-independent optimal key as the vector Φ that maximizes the objective function $\mathrm{d}\left(\Phi;\widehat{\mathbf{R}}\right)$ where $\widehat{\mathbf{R}}$ is the average correlation matrix of a speaker.

We construct a model for $\widehat{\mathbf{R}}$ for the voiced components of speech as follows. Let $\rho_l(k) = \frac{r_l(k)}{r_l(0)}$ be the normalized autocorrelation function (ACF) of the l^{th} frame. For each I, $\rho_l(k)$ has a peak at $k = 0$ and progressively smaller peaks at $k = k_p,\ 2k_p,\ \ldots$ where k_p is the pitch period which *is* assumed to be the same for all frames. On an average, the decay of the envelope of $\rho_l(|k|)$ may he assumed to he a linear function of k. Furthermore, $\rho_l(k)$ has significant values only for a few samples on either side of each peak. Hence a speaker-specific normalized average ACF is modeled as

$$\widehat{\rho}(k) = \begin{cases} \sum_{l=0}^{L-1}\rho_l\left(k - Ik_p\right)\left(1 - \frac{|Ik|}{N}\right), & |k - Ik_p| \leq M \\ 0 & \text{otherwise,} \end{cases} \qquad (8)$$

where M is typically between ***10*** and 15, and **I** takes integer values such that $|k| < \mathrm{N}$. This model is justified by the fact that, for small k, the average of $\rho_l(k)$ over several frames is known to have a small variance [11]. The (i,j) element of $\widehat{\mathrm{R}}$ is defined as $\widehat{\mathbf{R}}(i,j) := \widehat{\rho}(i-j)$.

For a given speaker, the frame-independent optimal key $\Phi_o(\widehat{\mathbf{R}})$ may be used to encrypt all the frames. For this key, the objective function $d\left(\Phi;\widehat{\mathbf{R}}\right)$ attains its maximum value $d_o(\widehat{\mathbf{R}}) := d\left(\Phi_o\left(\widehat{\mathrm{R}}\right);\widehat{\mathbf{R}}\right)$. However, the security of the encryption scheme may be increased by using a set of 6-suboptimal keys and switching the keys from frame to frame **in** a predetermined manner. For any 6 > 0, a sub-optimal key $\Phi_{so}\left(\widehat{\mathbf{R}}\ \delta\right)$ is defined as a vector Φ for which $\mathrm{d}\left(\Phi; 6\right) \geq d_o\left(\widehat{\mathbf{R}}\right) - 6$. For a given 6, let the set of all 6-suboptimal keys he described by S_δ.

Since the expression in Eq. (7) is a single equation with many variables, we have infinite solutions for any given value (except the maximum and minimum value) of $d_{MI}(\mathbf{x},\widetilde{\mathbf{x}})$. Hence S_δ is an infinite set. A random subset of S_δ may be selected, and keys belonging to this subset may **be** used in a predetermined order for encryption of successive frames of the signal. This strategy can enhance security and may **be** expected to yield a slightly small reduction in the distance value that provided by the optimal key. We note that any phase distortion of the optimally encrypted signal will destroy its optimality. Hence the effectiveness of the optimal key may be diminished if the nature of the encryption technique is known to the eavesdropper. A practical scenario may be like generating keys for diplomat A and diplomat B by making use of their voice characteristics.

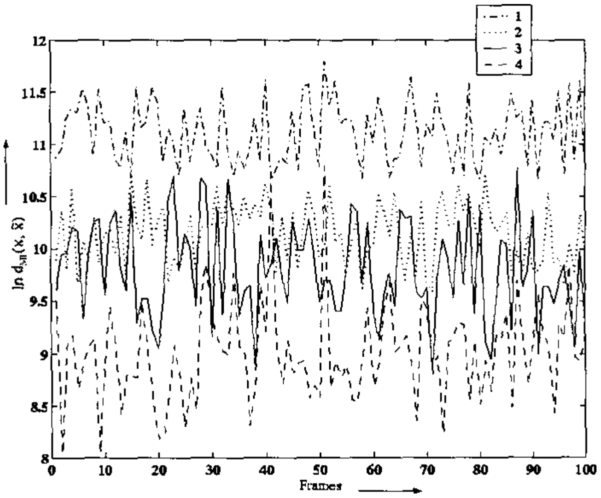

Fig. 3. The distances $\ln d_{MI}(\mathbf{x},\widetilde{\mathbf{x}})$ between the original and the encrypted signal frames with keys generated by different key generation schemes.

5. RESULTS

Experiments have been carried out to determine the performance of the encryption technique described above. The speech signal was sampled and divided into frames of length N = *512*. Four different key generation schemes were used for encryption. Results are presented in Figure 3 and Table 1 for 100 frames of speech of an adult female speaker whose average pitch was 269.3Hz. Only those frames whose energy exceeds the average energy per frame in the speech signal are included in this presentation. This selection was done to ensure that all the selected components frames had a component of voiced speech, since the model for the average correlation matrix $\widehat{\mathrm{R}}$ is valid only for voiced speech. The MI distance $\ln d_{MI}(\mathbf{x},\widetilde{\mathbf{x}})$ for 100 frames for the different key generation schemes are presented in Figure 3, while the average values of $\ln d_{MI}(\mathbf{x},\widetilde{\mathbf{x}})$ over 100 frames are given in Table 1. In scheme 1, represented by the dot-dash curve in Figure 3 each frame was encrypted using the optimal key $\Phi_o(\mathbf{R})$ corresponding to the actual correlation matrix **R** of that frame. The optimal key was obtained by finding the peak of $d\,(\mathrm{Q};\mathrm{R})$ by the conjugate gradient method for each frame. This method is not guaranteed to locate the global peak of $\mathrm{d}(\Phi;\mathrm{R})$. However, a comparison of the results for scheme 1with those for

Table 1. The average distance between the original and the encrypted signal frames with keys generated by different key generation schemes

Scheme	1	2	3	4
Averagedistance	11.11	10.09	9.87	***8.95***

scheme **4** (dashed curve in Figure 3), wherein the key for the encryption of each frame was generated randomly, indicates that the use of the optimal key has resulted in a large increase in the MI distance. Also, the frame-to-frame fluctuations in the MI distance are significantly reduced by the use of the optimal key. In scheme **2** (dotted curve in Figure 3), the key $\Phi_o(\widehat{\mathbf{R}})$ optimized for the average correlation matrix $\widehat{\mathbf{R}}$ was used to encrypt all the frames. The MI distances for this scheme are less than those for scheme 1, but significantly larger than those for scheme 4. The conjugate gradient method was used once again for determining $\Phi_o(\widehat{\mathbf{R}})$. With this key, the objective function $d\left(\mathbf{Q};\widehat{\mathbf{R}}\right)$ attains the peak value of 11.37. In scheme 3 (solid curve in Figure 3), a set of **20** $\delta-$suboptimal keys was used. Keys satisfying $d\left(\Phi;\widehat{\mathbf{R}}\right) \geq 11.0$ were obtained by selecting points around $\Phi_o(\widehat{\mathbf{R}})$ in **20** randomly chosen directions in the N-dimensional Q-space. This method for solving the roots involves the compuatation of the gradient and is refered to as the one-dimensional Newton's method in a multidimensional space [13]. Unlike the regular Newton's method, this root solving prcedure is not an iterative process. The gradient was evaluated using the mathematical package, Mathematica **[14]** that can do analytical differentiation. Keys from this set were then chosen randomly for encrypting the signal in each frame. A comparison of the results for schemes **2** and 3 indicates the MI distance values for scheme **3** are only slightly lower than for scheme **2.** It is clearly seen from Figure 3 that the $\delta-$suboptimal keys perform statistically better than the keys generated randomly (the solid curve Vs. the dashed curve) corroborating the theory and the heuristics presented for the distortion measure and in the estimation of the auto-correlation matrix, $\widehat{\mathbf{R}}$.

6. SUMMARY AND CONCLUSIONS

An analog encryption technique **that** is immune to a vector codebook based attack is proposed in this paper. Encryption is achieved by subjecting each frame of the sampled analog signal to a circulant transformation which introduces a phase distortion. A new distonion measure capable of quantifying phase distortion is also proposed. An optimal speaker-specific key generation scheme is then developed for maximizing an objective function based on the new distortion measure.Vulnerability to attack may be reduced by using a set of sub-optimal keys in a predetermined order instead of using a simple optimal key.

In the new distortion measure proposed by us the contributions of magnitude distortion and phase distortion are not segregated. Also, the relative imponance of magnitude distortion and phase distortion in speech encryption is not well understood. Hence, an objective comparison of the proposed encryption technique with other know techniques of analog encryption does not appear to he possible.

In the proposed encryption technique, the unintelligibility of the encrypted signal is due to a redistribution **of** energy throughout and within the expanse of the frame, and the consequent distortion of formants and introduction of new formants. Since a single frame of speech signal may contain acoustic cues of more than one word, the acoustic cues of the adjacent words mutually contribute to the unintelligibility of the encrypted speech. Hence a subjective test based on recognizing isolated words has not been considred and no other satisfactory subjective tests for comparing the performance of different encryption techniques has been developed. However, one could always combine the other scrambling techniques with the proposed encryption scheme i.e. perform a transposition of the DFT components (as in literature) apriori and then use the unitary circulant transformation. The scrambled signal would then be $\mathbf{AF^{-1}PFx}$, where P is the permutation.matrix used to trans-locate the DFT components. The analysis for key-generation for this kind of a combination of two methods can be done by considering the signal $\mathbf{F^{-1}PFx}$ as the original signal i.e. for a fixed permutation P , the circulant A may he obtained by the key generation scheme proposed.

7. REFERENCES

[1] A. Gersho and R. Steele, "Encryption of Analog Signals- A Perspective", IEEE 1. Select. Areas Commun., vol. SAC-2, No. 3, pp. 423-425, 1984.

[2] B. Goldburg , S. Sridharan and Ed. Dawson, "Design and Cryptanalysis of Transform-based Analog Speech Scramblers", IEEE J. Select. Areas Commun.. vol. 11, No. 5, pp. 735-743,1993.

[3] P. J. Davis, *Circulant Matrices*, New York. John Wiley and Sons, Inc., 1979.

[4] A.V. Oppenheim and J.S.Lim, "The Importance of Phase in Signals", Proc. IEEE, vol. 69, No.5, pp. 529-541, 1981.

[5] A. M. Libermann, F. S. Cooper, D. P. Shankweller, M. Studden-Kennedy, "Perception of the speech code", Psychological Review, vol. 74, No.6, pp. 431-461, Nov.1967.

[6] I. G. Proakis and D.G. Manolakis, ***Digital signal processing of signals: Principles, Algorithms and Applications,*** 3rd edition, Prentice Hall, 1996.

[7] G. Manjunath, Speech Encryption using Circulant Transformations, MSc(Engg.) thesis, 2001, Electrical Communication Department, Indian Institute of Science.

[8] S. R. Quackenbush, T. P. Barnwell and M. A. Clements, ***Objective measures of speech quality***, Prentice Hall, Englewood Cliffs, 1988.

[9] F. Itakura. "Minimum prediction residual principle applied to speech recognition", IEEE Trans. Acoust., Speech, Signal Processing. vol. ASSP-23, No.1, pp. 67-12. Feb. 1975.

[10] J. D. Markel and A. H. Gray, ***Linear prediction of speech,*** New York Springer Verlag, 1976.

[11] N. S. Jayant and P. Noll, ***Digital Coding of Waveforms,*** Prentice Hall, Englewood Cliffs, 1984.

[12] W. Rudin. ***Principles of Mathematical Analysis,*** Mc Graw-Hill Higher Education, 3rd edition, 1976.

[13] F. S. Action, ***Numerical methods that work,*** The Mathematical Association of America, Washington D.C., 1990.

[14] Wolfram Research, Inc., Mathematica, version 4, Champaign, IL, 1999.