International Journal of Engineering Research and Development
***e**-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com*
Volume 3, Issue 4 (August 2012), PP. 29-32

An Overview of Speech Encryption Techniques

Hemlata Kohad[1], Prof. V.R.Ingle[2], Dr.M.A.Gaikwad[3]
[1]*Electronics Engineering, R.C.E.R.T Chandrapur ,India*
[2,3]*Electronics Engineering,B.D.C.O.E.,Sevagram,India*

Abstract—Speech scrambling techniques are used to scramble clear speech into unintelligible signal in order to avoid eavesdropping. Analog scramblers are intended in applications where the degree of security is not too critical and hardware modifications are prohibitive due to its high cost. In this paper we discuss about the various techniques used for the encryption of the speech signal. PN sequences have random like properties, which helps in reducing the correlation among speech samples. The encrypted speech signal using different PN sequences is compared by informal listening tests and signal inspection methods in time and frequency domains.

Keywords—Scramblers, MSAAC, MSACC ,FOM.,intelligibility

I. INTRODUCTION

Speech encryption seeks to perform a completely reversible operation on a portion of speech, such that it is totally unintelligible to an unauthorized listener. The three most important criteria used to evaluate speech scramblers are

1) The scrambler's ability to produce encrypted speech with low residual intelligibility;
2) The extent to which the encryption and decryption processes affect the quality of the speech recovered by the intended recipient; and
3) The scrambler's immunity to cryptanalytic attack.

Cryptographers face the problem of designing scrambling systems which distort the very redundant speech signal to the extent that useful information is unable to be recovered .The encryption may be analog or digital .Analog speech encryption Also called speech scrambling we act on the speech samples themselves. In analog scrambling no modem or speech compression is required for transmission. The quality of the recovered speech is independent of the language .These scrambling schemes can easily be interfaced with the existing analog channels such as telephone, satellite or mobile communication links. Digital encryption involves digitization of the input speech signal first. The digitized signal is then compressed to produce a bit stream at a suitable bit rate this bit stream is encrypted and transmitted over the channel using modem. Digital encryption is more secure than analog ,but it needs complex implementation.

II. ANALOG SCRAMBLERS

This section briefly reviews the main concepts used in speech encryption techniques. They will be referred to as *Time-Segment Permutation* (TSP), *Frequency-Domain Scrambling* (FDS),*Time-Frequency Scrambling* (TFS), and *Transform-Domain Scrambling* (TDS).

A. Time-Domain Scramblers*:*

In time-domain scrambling [1], the most used method, TDS, divides the digitized speech signal x(n) into short time frames or blocks1 (typically 20ms, i.e., 160 samples for a sampling frequency fs = 8000kHz) which are divided in smaller segments that are permuted in time. Considering M time segments, **x**m , m = 1, . . .,M, each containing N samples, the i-th frame or block is represented as vector **x**i = $[\mathbf{x}_1^T\mathbf{x}_2^T\ldots\mathbf{x}_m^T]^T$with MN elements. An M×N matrix may be constructed with one segment per row, as follows:

$$\mathbf{X}i = [\mathbf{x}1\ \mathbf{x}2 \cdot \cdot \cdot \mathbf{x}M]^T \qquad (1)$$

The M×M permutation matrix **P** is defined as a matrix having only one nonzero element in each row, each nonzero element being equal to one. The scrambled speech block, vector **y**i, is obtained from the concatenation of the rows of the product **Y**i = **PX**i,

i.e.,_$\mathbf{y}i = [\mathbf{y}^T{}_1 y_2{}^T \ldots y_M{}^T]$ (2)
where
$\mathbf{Y}\ i = \mathbf{PX}i = [\mathbf{y}1\ \mathbf{y}2\cdot\cdot\cdot\ \mathbf{y}M]^T$ (3)

In the receiver, recovery of the original speech vector **x**i is obtained rearranging vector **y**i to form an M × N matrix **Y** i and multiplying it by matrix $\mathbf{P}^{-1}$.Time-domain scramblers have three major implementation factors that limit their application:
(i) need for time synchronization,(ii) introduction of time delay, and (iii) small effective number of keys [2]. Therefore, this class of scramblers has not been considered suitable for the application of interest in this work and was included here for didactic purposes only.

B. Frequency-Domain Scramblers*:*

Frequency-domain scramblers split the frequency contents of each speech signal block into M frequency bands. These bands are permuted according to some particular rule (or key), and a time sequence with scrambled frequency contents is synthesized to replace the original speech signal block. Frequency-domain scramblers are usually implemented with uniform filter banks or with wavelet transforms [3].An M-subband multirate filter bank is a set of M filters, which span the whole frequency spectrum. The speech signal is split into M subbands after passing through the analysis filter bank, Hi(z), and is critically downsampled, i.e., decimated by a factor of M. An M ×M permutation matrix **P** is inserted after the decimators in order to scramble the signals in the subband domain. It is then fed to upsamplers (interpolators) followed by the synthesis filter bank, Fi(z) (see Fig. 1).Recovery of the original speech vector is obtained with the same structure of Fig. 1 using the inverse of the permutation matrix, $\mathbf{P}^{-1}$.

C. Encryption by using Pseudo-Noise Sequences (EPNS):

Speech signal can be viewed as sequence of correlated samples and each sample as sequence of bits. The integibility of speech signal can be reduced by removing the correlation among the speech samples .PN sequence have random like properties which helps in reducing the correlation among speech samples. Speech encryption techniques using pseudo noise sequences make the speech signal unintelligible by removing the correlation between the samples of the speech signal .The sequences used for encryption in any case should not portray the statistical properties of the transmitted signal so that the attacker can not use statistical analysis to attack the system[4] .PN sequences [5]-[7] have noise like properties; these sequences are statistically independent and uniformly distributed. XOR operation of these sequences with the speech samples makes the speech signal a *noise-like* signal, and the encrypted speech signal sounds like a random noise signal. The PN sequences which have good auto-correlation and cross-correlation properties remove the correlation among the speech samples and results in an encrypted signal of less residual intelligence. The randomness of binary sequences is measured by mean square aperiodic auto-correlation (MSAAC) and mean square aperiodic cross-correlation (MSACC) measures [11]. The sequences which have better random noise properties will have less MSAAC and MSACC values.PN sequences are streams of 1's and 0's. Here, the waveform is taken as a *random-like,* meaning that it can be generated by mathematically precise rules, but statistically it satisfies the requirements of a truly random sequence in the limiting sense. These pseudo-random or pseudo-noise (PN) properties include, among other properties, (a) balance, (b) run and (c) auto-correlation properties [4]. These three properties make PN sequences efficient for speech encryption. However, due to the third property, adjacent bits correlation becomes considerably less, thereby making the PN sequences more effective for speech encryption when compared with data encryption due to high adjacent correlation present in the speech signals. Therefore, PN sequences that are useful for speech encryption must have very good auto-correlation and cross-correlation properties as well as maintaining some randomness properties. Below, we briefly describe different PN sequences useful for speech encryption. Note that in some cases we shall represent binary sequences using zeros and ones and in other cases +1's and ¡1's. The appropriate mapping is that the zeros are mapped to +1's and ones are mapped to -1's.

1. ***Maximal Length Sequences :***The Maximal length sequence (*m*-sequence) generator is usually constructed with linear feedback shift registers (LFSR) [11],[12]. The *m*-sequences are, by definition, the largest codes that can be generated by a given shift register of given length with feedback. The feedback function, also called as characteristic polynomial, determines the length and type of the sequence generated.
2. ***Gold Sequences:*** Gold sequences are generated by the modulo-2 operation of two different *m*-sequences of same length. Any two *m*- sequences are able to generate a family of many non-maximal product codes, but a preferred maximal sequences can only produce Gold codes [13].
3. ***Gold-Like Sequences*****:**There exists a class of sequences which has parameters similar to those of Gold sequences except that it is obtained from a decimated sequence. Let u be an *m*-sequence of length $N = 2n - 1$ generated by a primitive polynomial of degree n and let q be an integer such that $\gcd(q;N) = 3$. Also, let $v^{(k)}$, $k = 0; 1; 2$, denote the sequences obtained by decimating $T^k u$ by q. In that case, the new sequences formed by different combinations of u and v are called Gold-like sequences [13].
4. ***Barker Sequences :***Barker sequences are short length codes that offer good correlation properties. A Barker code is a sequence of some finite length N such that the absolute value of discrete autocorrelation function $|r(\tau)<= 1$ for $\tau = 0$ [14], [15]. Barker sequences have many advantages over other PN sequences. These sequences have uniformly low auto-correlation sidelobes (<=1), but the size of these families is small.
5. ***Barker-Like Sequences:***Barker sequences have good correlation properties with the peak correlation value being bounded by 1. The number of existing Barker sequences, however, are very less. We can generate more number of sequences by making certain relaxation on the peak value of the correlation function along with a maximum allowed shift between the sequences. This newly generated sequences are called Barker-like sequences [16].

6. ***Kasami Sequences:***Kasami sequences are also PN sequences of length $N = 2^n$-1, which are defined for even values of n [17],[18]. There are two classes of Kasami sequences: (i) small set of Kasami sequences, (ii) large set of Kasami sequences.Small set of Kasami sequences are optimal in the sense of matching Welch's lower bound for correlation functions. Asmall set of Kasami sequences [18] is a set of 2n=2 binary sequences. Fig. 2shows the block diagram representation for the generation of small set of Kasami sequences, each of length 63 bit. Small set of Kasami sequences are optimal sequences and have better correlation properties compared to Gold sequences. But the set contains less number of sequences. For the shift register of length n the number of possible sequences for the small Kasami sequence set is only 2$n/$2 sequences, whereas Gold code set contains 2^n + 2 sequences. The number of sequences can be increased by making some relaxation on the correlation values of the sequences. The resulting set of sequences is called large set of Kasami sequences [17],[18]. The performance of different PN sequences are evaluated by mean square aperiodic auto-correlation RAC (MSAAC) and mean square aperiodic cross-correlation RCC (MSACC) measures. Auto-correlation refers to the degree of correspondence between a sequence and phase shifted replica of itself,whereas cross-correlation is the measure of agreement between two different codes . If ci(n) represents non-delayed version of ck(i), cj(n+τ) represents the delayed version of ck(j) by 'τ ' units and N is the length of the sequence ci, then the discrete aperiodic correlation function is defined as

$$ri;j(\tau) = 1/N \sum_{\tau=1-N}^{N-1} ci(n)\, cj(n+\tau) \qquad (4)$$

The mean square aperiodic auto-correlation value for a code set containing M sequences is given by

$$R_{AC} = 1/M \sum_{i=1}^{M} \sum_{\tau=1-N}^{N-1} |\, ri;j(\tau)\,|^2 \qquad (5)$$

and a similar measure for the mean square aperiodic crosscorrelation value is given by

$$R_{CC} = 1/M \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \sum_{\tau=1-N}^{N-1} |\, ri;j(\tau)\,|^2 \qquad (6)$$

The sequences which have good auto-correlation properties will have poor cross- correlation properties and vice-versa, and they have wide and flat frequency spectrum. The sequences which have less MSAAC values removes the correlation among the bits within a sample, and the sequences which have less MSACC values removes the sample to sample correlation and make the speech signal less intelligible.

III. FIGURE OF MERIT

The price for being able to select good cross-correlation properties will be a degradation in auto-correlation properties of the set of sequences.A degradation of the auto – correlation properties has a direct relation on the frequency spectrum of the sequences in the set . If RAC values are poor the spectrum of the sequence not be wide band and flat.In order to determine quantitatively how significant this degradation is for given set of sequences , a FoM is required to judge the suitability of frequency characteristics of the sequences. Sequences with low FoM has narrow flat spectrum and they are neither suitable for CDMA nor for speech encryption.The FoM for a sequence ci(n) of length N having the auto-correlation function ri,j(τ) is given as :

$$Fx = r^{2}_{i,i}(0) / \sum |\, ri,j(\tau)\,|^2 \qquad \tau \neq 0$$

D. Combinations of different Encryption Methods*:*

Different types of encryption are presented. These are M=0:No encryption is required, M=1: Speech quality is high and security required is very less, M=2: Required security is low and power consumption and time delay are very less, M=3: Security required is moderate and the both time and power consumption are high, M=4: When the required security is higher than M=3 high power consumption and time delay, M=5: more secure than M=4, M=6: Security is higher with more time delay, M=7: When the transmitting party requires its voice stream to be secured highly and not considering the other effects (Power consumption and encryption delay).To provide more security, two levels or even three levels of encryption system can be employed according to the modes given in the Table I[10] Effectiveness of some of these modes is evaluated and is presented below.

E. Figure and Table

***Table I :* The different modes of encryption system**

Value Of M	**Encryption Method(s)**
0	*No Encryption is used*
1	*TPS*
2	*FDS*
3	*EPNS*
4	*TDS+FDS*
5	*TDS+EPNS*
6	*FDS+EPNS*
7	*TDS+FDS+EPNS*

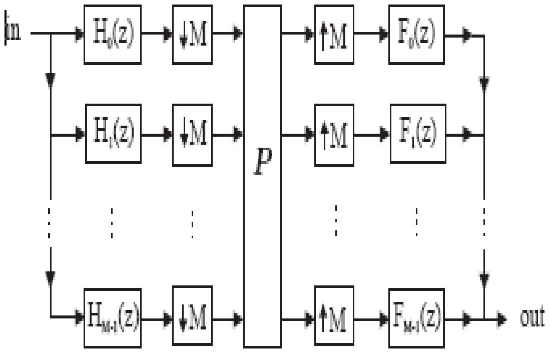

Fig 1 . Block diagram of frequency domain scrambler

IV. CONCLUSION

In this paper we have discussed different options available for the speech encryption .Depending upon the different parameters required for the encryption we can select the appropriate method .Speech encrypting using Kasami sequence is better with time domain and frequency domain as compared to the other PN sequences.Speech encryption security enhancement is possible usiing kasami sequence.

REFERENCES

[1]. N. Jayant, B. McDermott, S. Christensen, and A. Quinn, “A comparison of four methods for analog speech privacy,” IEEE Transactions on Communications, vol. COM-29, no. 1, pp. 18–23, January 1981

[2]. V. Senk, V. D. Delic, and V. S. Milosevic, “A new speech scrambling concept based on Hadamard matrices,” IEEE Signal Processing Letters,”vol. 4, no. 6, pp. 161–163, June 1997.

[3]. F. Ma, J. Cheng, and Y. Wang, “Wavelet transform-based analogue speech scrambling scheme,” Electronics Letters, vol. 32, no. 8, pp.719–721, April 1996.

[4]. Anil Kumar, Abhijit Mitra and S.R. Mahadeva Prasanna **“On** the Effectivity of Different Pseudo – Noise and Orthogonal sequences for Speech Encryption from Correlation Properties” International journal of information technology 2007

[5]. R. L. Pickholtz, D. L. Schilling and L. B. Milstein, “Theory of spread spectrum communications — A tutorial,” IEEE Trans. Commun., vol. COM-30, no. 5, May 1982.

[6]. E. H. Dinan and B. Jabbari, “Spreading codes for direct sequence CDMA and wideband CDMA cellular networks,” IEEE Commun. Magazine, vol. 36, no. 4, pp. 48-54, Sep. 1998.

[7]. B. Sklar, Digital Communications: Fundamentals and Applications, 2nd Ed., NJ: Prentice Hall, 2001

[8]. Dr. Nidaa A. Abbas, Member, IEEE “Speech Scrambling Based on PrincipalComponent Analysis” Manuscript received September 13, 2009.

[9]. Lin Shan Lee Ger-Chih Chou “A New Time Domain Speech Scrambling System Which Does Not Require Frame Synchronization” IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. SAC-2, NO, 3, MAY 1984 443

[10]. R. H. Laskar, F. A. Talukdar, B. Bora, K. S. P. Fernando, J. Anthony and L. Doley 9 “Complexity Reduced Multi-tier Perceptual Based Partial Encryption for Secure Speech Communication” 78–1–4244–4547–9/09/$26.00 c 2009 IEEE 1 TENCON EEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, VOL. 55, NO. 4, MAY 2008 1055

[11]. A. Fuster, L. J. Garcia, “An efficient algorithm to generate binary sequences for cryptographic purposes,” Theoretical Computer Science, vol. 259,pp. 679-688, May 2001.

[12]. S.Chattopadhyay,S.K.Sanyal,R.Nandi “Development of algorithm for the generation and correlation study of maximal length sequences for applicabilities in cdma mobile communication systems”

[13]. F. Rodríguez Henríquez,, N. Cruz Cortés ,, J.M. Rocha-Pérez . F. Amaro Sánchez. “Generation of gold-sequences with applications to spread spectrum systems”

[14]. S. W. Golomb and R. A. Scholtz, “Generalized Barker sequences,” IEEE Trans. Inform. Theory, vol. IT-11, no. 4, pp. 533-537, Oct. 1965.

[15]. D. G. Luenberger, “On Barker codes of even length,” Proc. IEEE, vol.51 pp. 230-231, Jan. 1963

[16]. C. K. Chan and W. H. Lam, “Generalised Barker-like PN sequences for quasisynchronous spread spectrum multiple access communication systems,” IEE Proc. Commun., vol. 142, no. 2, pp. 91-98, April 1995.

[17]. X. Wang, Y. Wu and B. Caron, “Transmitter identification using embedded pseudo random sequences,” IEEE Trans. Broadcasting, vol. 50, no.3, pp. 244-252, Sep. 2004.

[18]. D. V. Sarwate and M. B. Pursley, “Correlation properties of pseudorandom and related sequences,” Proc. IEEE, vol. 68, no. 5, pp. 593-619, May 1980.