A NEW SPEECH SCRAMBLING METHOD:
COMPARATIVE ANALYSIS AND A FAST ALGORITHM

V. D. Delić, V. Senk, and V. S. Milosević
University of Novi Sad, Faculty of Technical Sciences,
Trg Dositeja Obradovića 6, 21000 Novi Sad, Yugoslavia
Tel: (381 21) 350 244; fax: (381 21) 59 449
e-mail: tlk_delic@uns.ns.ac.yu

ABSTRACT

Conventional speech scrambling concept is based on permutation of time segments and/or frequency subbands. Although this approach is regarded as an insecure speech encryption method, almost all published scramblers are of that type. We found out that a linear combination based on Hadamard matrices instead of conventional permutation gives better cryptographic performances, maintaining all the good features of the scrambling concept. The new scrambling method provides a large keyspace and a simpler key selection. It attains negligible residual intelligibility and higher degree of cryptanalytic immunity. The price of these great improvements is a potential complexity increase. That is why we designed a fast algorithm for the new scrambling method.

1. INTRODUCTION

Most commercially available analog speech scrambling systems are based on permutation of speech components (SC) which are either time segments or frequency subbands or both [1]. Such a permutation hides the speech message content. However, it preserves the integrity of some essential features of the original waveform, preserving a good decrypted speech quality comparable to unprotected speech communication, but exhibiting insufficient security against cryptanalysis [2], [3]. Therefore, non-linear schemes were designed to increase security [4], [3]. However, there is an inherent problem with these conventional (permutation-based) scramblers: the deeper the scrambling the poorer the quality of the received and descrambled signals, while the implementation complexity increases.

In this paper, we introduce a new speech scrambling concept [5]. It is based on Hadamard matrices instead of conventional permutation, providing better cryptographic performances. Moreover, the new speech encryption method preserves both the bandwidth and the speech quality. A fast algorithm for the procedure is also designed in order to preserve low complexity and price - important features of analog scramblers.

2. SCRAMBLING AS LINEAR TRANSFORM

Conventional permutation scramblers and our Hadamard matrix-based procedures can both be represented as liner transforms given by

$$y = x \cdot [S]; \qquad (1)$$

where x and y are row vectors of N plain and N scrambled SC, respectively, and $[S]$ is the scrambling matrix.

If delay is limited, than not more than $N = 8$ or 16 time segments are used. A low number of frequency subbands ($N = 2$ or 4) is used due to filter non-linearity, except for the so-called transform-based scramblers where subbands are transform coefficients; N is much greater then.

2.1. Scrambling based on permutation

If the scrambling is based on permutation, the scrambling matrix has got one and only one non-zero number, equal to 1, in each row and column. Their coordinates define the scrambling key. There are $N!$ keys, but only 10–20% of them provide a sufficiently low residual intelligibility. That is why the key selection is a complex request, especially for great N.

The permutation matrix has some features necessary for speech scrambling/descrambling. It is linear, orthogonal and normalised. The inverse matrix $[S]^{-1}$ which is used for descrambling is equal to $[S]^T$. It always exists and does not enhance the channel introduced deformations. These features are the reason why the permutation-based scrambler preserves the essential speech signal features such as bandwidth and dynamics and thus the speech quality.

However, the cryptanalysis of speech scramblers based on the time segment permutation is possible due to relatively slow spectrum change at the segment boundaries. The cryptanalysis of conventional frequency scramblers (including transform-based ones) is even more efficient, and based on code books with vectors of typical spectral parameters [2], [3].

There is a number of techniques which improve the scrambling performances. For instance, simple time and frequency reversion, sliding window method [1], and the so-called two-dimensional scrambling, i.e. simultaneous time and frequency SC permutation [6]. Non-linear dummy component insertion is also possible [4], but the permuting procedure cannot hide these dummy components effectively.

2.2. Scrambling based on Hadamard matrices

In the new scrambling concept, speech components are linearly combined contrasted to their conventional permutation. The linear combination is one of H-equivalent normalised Hadamard matrices [7].

An N-dimensional Hadamard matrix is an N by N matrix $[H_N]$ of 1's and −1's with

$$[H_N] \cdot [H_N]^{t} = N \cdot [I_N]. \qquad (2)$$

The dimension N must be 1, 2, or 4n, $n \in Z^+$. Hadamard matrices may be transformed into other Hadamard matrices by signed permutations of rows and columns. Such matrices are called H-equivalent. A Hadamard matrix is H-normalised if its first row and column consist entirely of 1's. For instance, $[H_1] = [1]$ and

$$[H_2] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \qquad (3)$$

There exist two simple constructions of Hadamard matrices. The first one is Silvester's and is given by $[H_{N_1 N_2}] = [H_{N_1}] \otimes [H_{N_2}]$, where $\otimes$ is the Kronecker product. Using $[H_2]$ matrix given above, this construction produces Hadamard matrices for all $N = 2^n; n \in Z^+$. The second construction is Paley's one. It gives Hadamard matrices of order $N = p^n + 1; n \in Z^+$ and p is a prime. A set of other less general constructions is given in [7].

The matrix [S] in (1) is now chosen to be the scaled version of one of H-equivalent Hadamard matrices, obtained using

$$[S] = \frac{1}{\sqrt{N}} [P_r][H_N][P_c]; \qquad (4)$$

where $[P_r]$ is the row and $[P_c]$ the column signed permutation matrix, both having only one non-zero element in each row and column, equal to either +1 or −1, and [H] is the original (usually H-normalised) Hadamard matrix.

Notice that such a scrambling matrix is also orthogonal and normalised, but each scrambled component depends on all the plain ones, i.e. the speech components are altered themselves. For $N = 2$ and $[P_r] = [P_c] = [I_2]$, this procedure resembles the one used for stereo transmission.

3. COMPARATIVE ANALYSIS OF THE NEW AND CONVENTIONAL SCRAMBLERS

3.1. Speech quality

Because of the features of the descrambling matrix $[S]^{-1}$ previously discussed, the new scrambler preserves the speech quality. We experimentally tested both time and frequency scramblers, permutation scrambler as well as the new one. For different SNR we observed noise power before and after descrambling.

Table 1: Noise power before and after descrambling
(t-time, f-frequency, p-permutation, h-Hadamard)

σ^2 (SNR)	t.p	t.h	f.p	f.h
106530 (10)	106488	107310	106489	106491
10653 (20)	10730	10783	10731	10729
1065 (30)	1079	1073	1079	1079
106.5 (40)	108.0	106.8	107.6	107.8
10.6 (50)	11.3	11.0	11.1	11.2

Channel introduced deformations are not amplified during descrambling.

3.2. Residual intelligibility

Some experiments have been carried out in order to compare the residual intelligibility of the new speech scrambler with the conventional ones. About 50 listeners (students) tried to recognize 1 of 10 known sentences that lasted 5 seconds. Eight segments were scrambled. First, they lasted 32ms ('), then, some of them were inverted ("), and finally, they lasted 64ms ("') (no inversion here). Simple guessing gives a recognition rate of 10%. Black color corresponds to the conventional permutation-based scrambler, and lighter color to the new one.

The results in Figure 1 show that the new scrambler achieves desirable residual intelligibility - near to zero. Notice that time inversion is more useful than using longer blocks in permutation-based scramblers, because segments remain unaltered. The Hadamard-based scrambler alters the signal inside segments. So scrambling using longer blocks gives lower residual intelligibility. Notice also that consonants are being hidden beneath vowels during their linear combination, because vowels have greater power. The cryptogram sounds as singing vowels. All the tested keys provide approximately the same - negligible residual intelligibility. Thus, there is no need to search for good keys in the new scramblers.

3.3. Keyspace

Conventional permutation-based scrambling concept has keyspace with just about 10% of N! keys that can provide sufficiently low residual intelligibility. It is not the

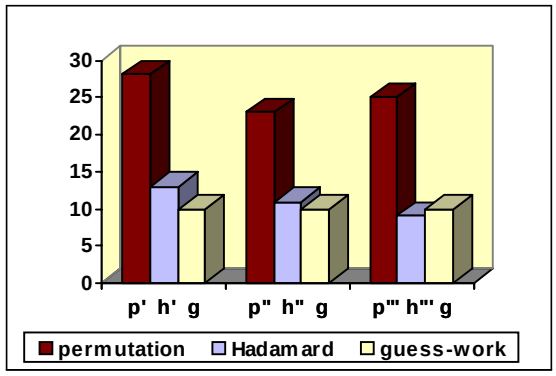

Figure 1: Recognition of 1 Among 10 Sentences

reduced number of keys, though, that imposes the greatest problem, but the complicated verification of key suitability for use (the key selection problem).

There are $N!2^N$ H-equivalent Hadamard matrices of order N, but some of them are identical. Namely, all the combinations of signed permutation of rows and columns which transform a Hadamard matrix $[H_N]$ into itself make the automorphism group $A([H_N])$. Thus, the number of non-identical H-equivalent matrices of the $[H_N]$ matrix is given by

$$\nu = \frac{N!2^N}{\#A([H_N])}; \tag{5}$$

where $\#A([H_N])$ is the cardinality of the automorphism group. The new concept provides a keyspace with

$$\Psi = \sum_{i=1}^{\rho_N} \nu_i \tag{6}$$

keys, where ρ_N is the number of H-inequivalent Hadamard matrices of order N, and ν_i are given by (5) (the cardinality of automorphism groups of H-inequivalent Hadamard matrices usually differ). This number is much greater than $N!$.

Table 2: Number of bits per key
(' = counted, " = estimated)

N	8	12	16
permutation	12	26	42
Hadamard	32'	64'	100"

For instance, DES is a well-known data encryption standard whose key has 56bits (each new bit doubles the number of keys). In Table 2 we can see that the permutation based scrambler cannot attain more than 42 bits, even with 16 segments, and we know that the delay is critical. But the new speech scrambler already exceeds DES with 12 speech components, and with 16 segments has a really large number of keys, providing sufficient protection against brute force code-breaking attacks.

It should be noted that it is not the key but the speech statistics that is attacked in the linear scrambling case, since there is residual redundancy in the cryptogram. The new speech scrambler, altering the speech components themselves, effectively scrambles the kind of speech statistics that is presently used for cryptanalysis of conventional speech scramblers. Moreover, the cryptogram sounds much the same with all the possible keys, a characteristics that eliminates the need for key selection.

4. FAST ALGORITHM

The price for these great improvements is a potential complexity increase. Using decomposition of a Hadamard matrix, a fast algorithm for multiplying with $[H]$ may be designed. The complexity of the procedure is thus reduced from N^2 in brute force implementation to M, where $N^2 \gg M$, $N \log N$.

We have designed three types of fast algorithms. The first one uses a simple decomposition of Sylvester type Hadamard matrices as

$$[H_N] = [A]^{\log_2 N} = [A^T]^{\log_2 N}; \tag{7}$$

the so-called Good decomposition. For instance, for $N = 8$, the matrix $[A]$ has the form

$$[A] = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}; \tag{8}$$

and requires N addition/subtraction operations.

The second type of fast algorithm also makes decomposition of Silvester type Hadamard matrices, but using the Kronecker product representation given by

$$[H_{2N}] = [H_2] \otimes [H_N] = \begin{bmatrix} [I_N] & [I_N] \\ [I_N] & -[I_N] \end{bmatrix} \cdot \begin{bmatrix} [H_N] & [0_N] \\ [0_N] & [H_N] \end{bmatrix}; \tag{9}$$

In the same example, such a decomposition gives

$$\begin{aligned} [H_8] &= [H_2] \otimes [H_2] \otimes [H_2] \\ &= [A_1] \cdot [A_2] \cdot [A_3]; \end{aligned} \tag{10}$$

where

$$\begin{aligned} [A_1] &= [H_2] \otimes [I_4]; \\ [A_2] &= [I_2] \otimes ([H_2] \otimes [I_2]); \\ [A_3] &= [I_4] \otimes [H_2]: \end{aligned} \tag{11}$$

If the Hadamard matrix is not of Silvester type, say for $N = 12$, than the decomposition can be made along the lines of the Cooley-Tukey algorithm, designed for FFT [8], and modified for FHT [9]. The signal flow diagrams for Sylvester type $N = 8$ Hadamard matrix and the one for $N = 12$ are shown in Figure 2.

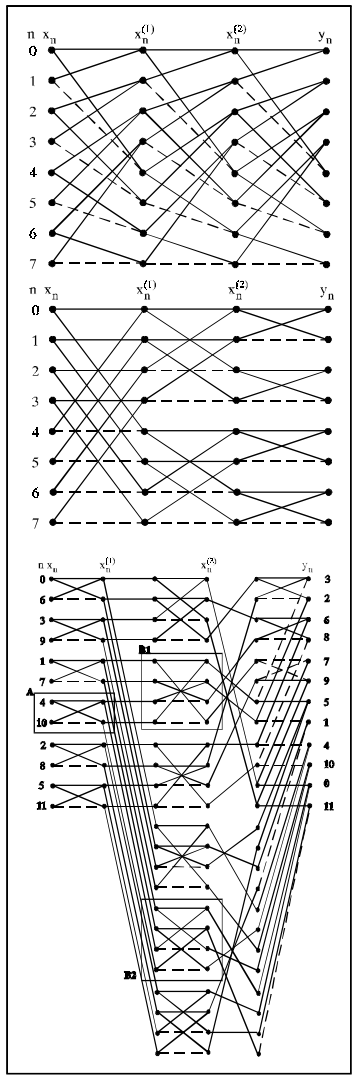

Figure 2: Signal Flow Diagrams for Fast Hadamard Transform (N = 8 Good, N = 8 Cooley–Tukey, N = 12 Cooley Tukey type; - - - denotes multiplication by -1)

5. CONCLUSION

The new scrambling concept is a speech encryption method which combines many good features of both digital and analog methods. First, it has a large keyspace and a simple key selection like the digital speech encryption systems. Then, the new scrambling concept provides a negligible residual intelligibility. Moreover, it scrambles the kind of speech statistics presently used for cryptanalysis of conventional speech scramblers, providing a higher degree of security. On the other hand, it maintains the signal bandwidth and speech quality as well as low system complexity (the fast algorithm) and cost, i.e. all the good features of permutation scrambling approach. Also, it better hides inserted dummy components than conventional scramblers do. Finally, this new scrambling concept is compatible with conventional ones.

References

[1] H. J. Beker and F. C. Piper, Secure speech communications, London, UK, Academic press, 1985.

[2] E. Dawson, B. Goldburg, and S. Sridharan, The Automated Cryptanalysis of Analog Speech Scramblers, Eurocrypt'91 Abstracts, pp. 203–207, Brighton, UK, April 1991.

[3] B. Goldburg, S. Sridharan, and E. Dawson, Design and Cryptanalysis of Transform-Based Analog Speech Scramblers, IEEE Journal on Selected Areas in Communications, Vol. 11, No. 5, pp. 735–744, June 1993.

[4] A. Matsunaga, K. Koga, and M. Ohkawa, An Analog Speech Scrambling System Using the FFT Technique with High-Level Security, IEEE Journal on Selected Areas in Communications, Vol. 7, No. 4, pp. 540–547, May 1989.

[5] V. Senk, V. D. Delić, and V. S. Milosević, A New Speech Scrambling Concept Based on Hadamard Matrices, IEEE Signal Processing Letters. (under consideration)

[6] N. S. Jayant, R. V. Cox, B. J. McDermott, and A. M. S. Quinn, Analog Scramblers for Speech Based on Sequential Permutations in Time and Frequency, The Bell System Technical Journal, Vol. 62, pp. 25–46, January 1983.

[7] W. D. Wallis, A. P. Street, and J. S. Wallis, Combinatorics: Room Squares, Sum-Free Sets, Hadamard Matrices, Springer-Verlag, Berlin, 1972.

[8] R. E. Blahut, Fast Algorithms for Digital Signal Processing, Addison-Wesley Publishing Company, New York, April 1985.

[9] I. Radujkov, V. Senk, and V. D. Delić, An Algorithm for Fast Hadamard Transform, Proceedings of National Conference on ETRAN, Budva, YU, June 1996.