CRYPTANALYSIS OF SPEECH SIGNALS CIPHERED BY TSP USING ANNEALED HOPFIELD NEURAL NETWORK AND GENETIC ALGORITHMS

J. A. Apolinário Jr.[1,3] apolin@coe.ufrj.br
P. R. S. Mendonça.[1] mendonca@coe.ufrj.br
R. O. Chaves.[1] chaves@coe.ufrj.br
L. P. Calôba.[1,2] caloba@coe.ufrj.br

[1] COPPE / UFRJ, Brazil
[2] EE / UFRJ, Brazil
[3] Instituto Militar de Engenharia, Brazil

ABSTRACT

This paper deals with the cryptanalysis of speech signals ciphered by Time Segment Permutation. It shows that (a) if the number of segments-per-block is not large, the distance between the cepstral coefficients of the segments edges may be used as a vicinity criteria and (b) the combinatorial optimization problem of re-ordering the segments may be solved using a Neural Network improved with Simulated Annealing or a Genetic Algorithm. Experimental results are presented.

1. INTRODUCTION

THE cryptanalysis of speech signals ciphered by time segment permutation (TSP) of fixed size and block to block (jumping window) can be successfully done by exhaustive search in the space of all possible permutations when the number of segments is small enough, e.g. 8 segments [1]. The basic idea is to choose the permutation that results in the smallest sum of spectral distances between edges of adjacent segments. This work presents three different methods to solve this problem: Hopfield Neural Networks associated with Simulated Annealing and Mean-Field Annealing and Genetic Algorithms.

The scheme of cryptanalysis will be briefly presented in Section 2. The formulation of the problem with a Hopfield Network can be found in Section 3. In Section 4, Simulated Annealing is introduced as an attempt to improve the quality of the results. Section 5 brings the Genetic Algorithms in order to decrease the processing time. Section 6 presents the experimental results and, finally, Section 7 summarizes some conclusions.

2. CRYPTANALYSIS SCHEME

Fig. 1 shows the concepts of blocks and segments. Since the segments will be permuted within a block, they must be stored in memory before transmission, resulting a total delay of twice the block size (transmission and reception). This delay is one of the limitations of the process as well as the number of segments on each block: a large number of segments would decrease the residual intelligibility and increase the resistance to cryptanalysis, but at the same time would cause bandwidth expansion, need of a more precise synchronism and enhancement in the overlapping of segments when the signal is transmitted through a channel.

In this work the measure of similarity between the end and the beginning of two speech segments is done by computing how close are the spectra of the end of a segment and the beginning of another one.

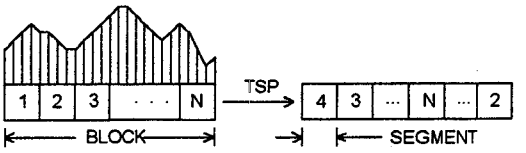

Fig. 1 - Speech signal divided in blocks and segments.

It can be found in literature [1] several measures of spectral distance. All of them are widely used in speech processing and particularly in scramblers [9] to objectively measure the residual intelligibility and the quality of the recovered voice. This work uses the cepstral coefficients [7] to define a cepstral distance given by Eq. 1, which is associated to the mean square difference of two logarithm magnitude spectra:

$$cd(R,L) = \left[c_R(0)-c_L(0)\right]^2 + 2\sum_{i=1}^{p}\left[c_R(i)-c_L(i)\right]^2 \quad (1)$$

where c's are the cepstral coefficients in R (from right) and L (from left), and p is the order of the model. We observe in Eq. 1 that the information of energy is already included (c's of zero).

Due to the need of computing the coefficients of the speech signal model in a short time analysis, we use a lattice structure that, besides offering the possibility of obtaining the coefficients adaptively, presents coefficients with absolute values less than one and allows calculations directly from the samples without intermediary computation of the autocorrelation function [3].

The lattice structure coefficients, often called reflection coefficients $\{k_i\}$, do not depend on filter order and can be transformed into $\{a_i\}$ (LPC) or $\{c_i\}$ (cepstral) coefficients. The algorithm used to estimate the coefficients was the SQNLSL (Square-Root-Normalized Least-Square Lattice). This algorithm has lower computational complexity when compared to the RLSL (Recursive Least-Square Lattice), and better numerical properties [3].

The proposed scheme can be applied on a speech signal ciphered by the permutation of N time segments with identical size and confined into blocks transmitted sequentially. It will be assumed here that N = 8 in order to compare with results obtained by exhaustive search. The main reason for using algorithms other then the exhaustive search is the possibility of extending their use to the case of a higher number of segments.

0-7803-3636-4/97 $10.00 © 1997 IEEE

The algorithm presented in Fig. 2 shows how the reordering of the segments will be processed. The search for a minimum distance permutation is carried out by the methods of sections 4, 5 and 6.

```
/* CRYPTANALYSIS ALGORITHM */
BEGIN TSP2VOICE (INPUT: SIGNAL.TSP, OUTPUT: SIGNAL.VCE)
  "PROGRAM INITIALIZATION";
  WHILE NOT END OF FILE
    "READ A BLOCK FROM INPUT FILE":
    FOR SEG FROM 1 TO N STEP 1
      "CALCULATE K_R (RIGHT OF EACH SEGMENT)";
      "CALCULATE K_L (LEFT OF EACH SEGMENT)";
    END-FOR;
    "CALCULATE CEPSTRAL COEFFICIENTS FROM THE {k_i}";
    "CALCULATE DIST. d[i][j] (RIGHT SEG i TO LEFT SEG j)";
    "FIND THE LEAST DISTANCE PERMUTATION";
    "STORE THE COEFFICIENTS OF THE FINAL BLOCK";
    "REORDER THE SEGMENTS";
  END-WHILE;
END {TSP2VOICE}.
```

Fig. 2 - Cryptanalysis algorithm (BLOCK with N segments).

3. STATING THE PROBLEM

Hopfield Networks has been widely used as a tool in the search of solutions to combinatorial optimization problems. It can be proved [6] that a Hopfield Network with discrete output neurons operated in an appropriate manner is a stable system in the Lyapunov sense. The Lyapunov function used in this work is given by

$$E=-\frac{1}{2}\sum_{\substack{ijkl\\ k\neq i\\ l\neq j}} w_{ijkl}y_{ij}y_{kl}-\frac{1}{2}\sum_{\substack{ijk\\ k\neq i}} w_{ijkj}y_{ij}y_{kj}-\frac{1}{2}\sum_{\substack{ijl\\ l\neq j}} w_{ijil}y_{ij}y_{il}-\sum_{\substack{ij\\ i\neq j}} w_{ijij}y_{ij}^2-\sum_{ij} t_{ij00}y_{ij}+ct \quad (2)$$

where y_{ij} corresponds to neuron ij output, t_{ij00} corresponds to the bias applied on neuron ij and w_{ijkl} corresponds to the synapses connecting neuron ij output to neuron kl input.

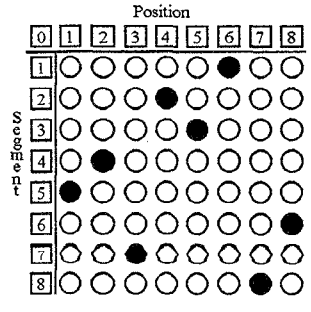

Fig. 4. Each circle corresponds to a neuron and the filled circles show the activated neurons. The permutation of the above configuration is 5-4-7-2-3-1-8-6.

Thus, it is necessary to state the cryptanalysis problem in such a way that its solution is equivalent to the minimization of an objective function with the form of Eq. 2. This problem is very similar to the well known Traveling Salesman Problem. A proposal to such function is $F = F_A + F_B + F_C + F_D + F_E$, where,

$$F_A = \sum_{ij} A_{ij}y_{ij}\left(1-y_{ij}\right), \quad (3)$$

(minimal when $y_{ij} \in \{0, 1\}$)

$$F_B = \sum_i B_i\left(1-\sum_j y_{ij}\right)^2, \quad (4)$$

(minimal when only one active neuron is on each column)

$$F_C = \sum_j C_j\left(1-\sum_i y_{ij}\right)^2, \quad (5)$$

(minimal when only one active neuron is on each row)

$$F_D = \sum_{\substack{ijk\\ i\neq k}} D_{ijk}y_{ij}\left(d_{ki}y_{k,j-1}+d_{ik}y_{k,j+1}\right) \quad (6)$$

(assuming other terms already minimized, F_D becomes minimal when the permutation found yields the least distance)

and

$$F_E = \frac{1}{2}\sum_i E_i d_{0i}y_{i1} \quad (7)$$

(similar to the above, except for referring to the distance between the last segment of the previous block and the first segment of the present block).

Stating that $A_{ij} = A$, $B_i = B$, $C_j = C$ and $D_{ijk} = E_i = D\ \forall$ (i, j, k), rewriting F and comparing it term to term with E (Lyapunov function) in Eq. 2, we obtain the weights of the synapses of the desired Hopfield Network, which are shown next.

$$t_{ij00} = \begin{cases} -A+2(B+C), & \text{if } j \neq 1; \\ -A+2(B+C)-(1/2)\mathrm{Dd}_{0i}, & \text{if } j = 1; \end{cases}$$

$$w_{ijij} = A-B-C;$$

$$w_{ijil} = -2B; \qquad w_{ijkj} = -2C;$$

$$w_{ijkl} = \begin{cases} Dd_{ki}, & \text{if } j = l+1; \\ Dd_{ik}, & \text{if } j = l-1; \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

The next step is to adjust the values of A, B, C and D. Since B and C are the terms responsible for keeping one neuron per row and one neuron per column in the diagram of Fig. 4, it is reasonable that they have the same value resulting in the same penalty associated to the violation of these two conditions. After several experiments, it was verified that the best value of A is the one that nulls the weights of the self-feedback synapses, given by w_{ijij}. So, we found $A = B + C$, and adopted $B = C = 1$ and $A = 2$.

The adjustment of coefficient D was a difficult task. Small values could result in a low cost to a permutation corresponding

to a high distance, while large values could result in a low cost to the violation of the validity condition F_B and F_C, generating a solution that does not have only one active neuron per row or per column. So, we decided to introduce an adaptive scheme to set D: the network is operated initially with D equal to the inverse of the mean value of the cepstral distances between all its segments. When the solution found is not valid, we decrease D according to the equation $D_{new} = (1 - \alpha_{down}) D_{old}$. If the solution found is valid, D is increased by the equation $D_{new} = (1 + \alpha_{up}) D_{old}$. The parameters α_{up} and α_{down} can be, in general, equal, and their best values were fixed in 0.5.

For the sake of comparison, another solution was generated. The last segment of the previous block was followed by the nearest segment of the present block, and so forth, until all segments were used. This solution was called *the nearest segment solution.*

4. SIMULATED ANNEALING

The Hopfield Network as a combinatorial optimization method usually fails in solving non-convex problems due to the presence of local minima, which cause a premature stop of the algorithm. Techniques based on Simulated Annealing, Mean-Field Annealing and Genetic Algorithms [2] are possible solutions to this problem. The first two are based on the phenomenon of metal annealing, from where it is known that if the cooling of a metal is slow enough, its atoms will reach a minimum energy configuration, despite they can make, with a certain probability, temporary transitions to states of higher energy. So, the idea is to disturb the network by the injection of random noise (equivalent to heating), and slowly decrease the variance of the noise (cooling scheme), in an attempt to obtain a network converge to the global minimum of its Lyapunov function.

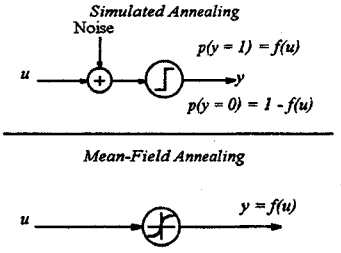

Fig. 6. Comparison between neurons used in the techniques of (a) Simulated Annealing and (b) Mean-Field Annealing.

In the technique of Simulated Annealing, we add uncorrelated random noise with zero mean and a given distribution to each neuron's input signal of the Hopfield Network (Fig. 5a). This means that the neuron output will not be a deterministic function its the input signal, but a stochastic process with distributions determined by the input signal and the noise statistics. It can be proved [10] that the use of a Cauchy distribution allows a fast cooling scheme, with a hyperbolic temperature decay. In the technique of Mean-Field Annealing (Fig. 5b) the discrete neuron is substituted by a continuous one, which, when presented to an input signal, gives as output the mean value of a discrete neuron output to whose input was added a random noise.

Our experiments suggested that when a Cauchy distribution is taken the use of Simulated Annealing produced better outcomes. Another feature used was the conditional acceptance of a transition that yields a growth in the cost function with probability given by $e^{-\frac{\Delta E}{T}}$, where ΔE is the increase in the cost function and T is the temperature of the network, corresponding to the variance of the noise added to the input.

Combining the techniques of Simulated Annealing and adaptive adjustment of D, the algorithm found solutions which are global optimal (the same solution of the exhaustive search) in 100% of the our experiments (8 segments TSP scramblers).

5. GENETIC ALGORITHMS

Genetic Algorithms (GA's) [4,5] can be defined as computational methods of search based on the mechanisms of natural selection and genetics. In GA's there are always one or more groups, called "populations", that are composed of possible solutions of a given problem. This population evolves according to probabilistic operators ingenious on biological metaphors with the purpose of, in average, make the individuals represent improvement of the solutions as a consequence of the evolutionary process.

The evolutionary process of a population can be characterized by the crossing of its individuals, expecting an improvement of them. For most applications of GA's, the population size is constant. At each crossing, new individuals are created. The renovation process of the population is given by the natural selection, that attempts to preserve the most adapted individuals and eliminates the fewer ones. The survival probability of each individual is proportional to its adaptation. Events like recombination, mutation and extermination can represent some examples of strategies that make the evolutionary process even more dynamic. According to the features of a given problem, the use of these events can improve the performance of the GA.

To the cryptanalysis problem it was used a population with constant size of 100 individuals. In the initialization process of the population, each individual was generated at random, in order to achieve a good dispersion of them on the search space. At the crossing stage, each pair of individuals was selected at random to generate new others. The survival probability of them grows with its adequability function (adaptation level). The methodology used to achieve the crossing was the ERX method [8], "Edge Recombination Crossover", that is considered very effective to transfer common characteristics between a given pair of individuals. The natural selection strategy can be regarded elitist, because the best individual is always preserved.

Among the algorithms presented in the last sections, GA's show the least processing time. The overall performance was very good, since for each ciphered signal the result was always the same obtained by the exhaustive search method.

6. EXPERIMENTAL RESULTS

The results of the simulations are presented by objective measures of the similarities between the cryptanalysed and the

original signals. It might be noted that the objective here is not the perfect reconstruction of the speech signal, but only to obtain the intelligibility of the ciphered messages.

The sample signals used in the tests were generated by a "jumping window" TSP scrambler simulator with 8 segments of 256 samples per block. The following speech files were digitized with a sample frequency of 8KHz, 12 bits per sample:

- school.vce: male speaker, English;
- sinto.vce : female speaker, Portuguese, recorded from a TV channel;
- telef.vce: male speaker, Portuguese, from a telephonic channel;
- vega.vce: female speaker, capela music in English.

We can observe in Tab. 1 a measure of relative spectral distortion for each cryptanalysed signal, when compared with the original signals. This distortion corresponds to the Euclidean distance between the $\{a_i\}$ (LPC) coefficients estimated segment to segment of the original and cryptanalysed signals over the norm of the coefficients of the original signal.

Another objective performance measure is the ***hit rate*** (number of segments positioned in their proper places over the total number of segments), shown on Tab. 2.

	SCHOOL	SINTO	TELEF	VEGA
NEAREST SEGMENT SOLUTION	41.2%	92.1%	39.6%	68.4%
MEAN FIELD ANNEALING	41.0%	39.1%	34.7%	40.7%
SIM. ANNEAL. WITH ADAPTIVE *D* AND GA'S	23.3%	24.0%	21.3%	5.3%
EXHAUSTIVE SEARCH SOLUTION	23.3%	24.0%	21.3%	5.3%

Tab. 1 - Relative spectral distortion.

	SCHOOL	SINTO	TELEF	VEGA
NEAREST SEGMENT SOLUTION	52.3%	12.5%	22.5%	19.8%
MEAN FIELD ANNEALING	56.2%	40.6%	32.5%	56.6%
SIM. ANNEAL. WITH ADAPTIVE *D* AND GA'S	72.7%	65.6%	63.1%	91.2%
EXHAUSTIVE SEARCH SOLUTION	72.7%	65.6%	63.1%	91.2%

Tab. 2 - Hit rate (# correct segments / # segments).

7. LAST ANALYSIS AND CONCLUSIONS

From the results obtained in the Section 6, we conclude that the proposed methods (Simulated Annealing and GA's) succeeded in recovering the intelligibility of the ciphered signal, achieving a result identical to that obtained with the use of an exhaustive search. The great advantage that arises at this point is the possibility of using the method in the cryptanalysis of TSP scramblers with a higher number of segments. Initial investigations with a 16 segments scramblers pointed the Genetic Algorithm as the prospective victorious method due to its reasonable amount of processing time in finding the possible best key among a search space of 2×10^{13} solutions. Nevertheless the cause of the unbearable results obtained with 16 segments seems not to be the optimization technique but the spectral distance used since the original signal has a "distance" much higher then the solution of the GA scheme. This fact indicates that a better measure of vicinity between segments is needed.

REFERENCES

[1] Apolinário Jr., J. A., Criptoanálise de Sinais de Voz Cifrados por Permutação de Segmentos Temporais, MS Thesis, UnB, Brasilia, June/1993.

[2] Cichoki, A. and Unbehauen, R., Neural Networks for Optimization and Signal Processing, John Willey, 1993.

[3] Cowan, C.F.N. and Grant, P.M., Adaptive Filters, Englewood Cliffs, Prentice-Hall, 1985.

[4] Goldberg, D.E., Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, 1989.

[5] Holland, J.H., Adaptation in Natural and Artificial Systems, Univ. of Michigan Press, 1975.

[6] Hopfield, J.J., Neural networks and physical systems with emergent collective computational abilities, in "Proceedings of the National Academy of Sciences - USA", vol. 79, April/1982, pp. 2554-2558.

[7] Markel, J.D. and Gray Jr., A.H., Linear Prediction of Speech, Berlin, Springer-Verlag, 1976.

[8] Michalewicz, Z., Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag, 1992.

[9] Sridharan, S., Dawson, E. and Goldburg, B., Fast Fourier Transform Based Speech Encryption System, in "IEE Proceedings-I", vol. 138, n°. 3, June/1991, pp. 215-223.

[10] Szu, H. and Hartley, R., Fast Simulated Annealing, Physical Letters, vol. 122, n°.3, 4, June/1987, pp. 157-162.

[11] Wasserman, Phillip D., Neural Computing - Theory and Practice, Van Nostrand Reinhold, 1989.