Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **2**, March **2016**

An Unconditionally Secure Speech Scrambling Scheme Based on An Imitation Process to A Gaussian Noise Signal

Dora M. Ballesteros L., Diego Renza and Steven Camacho

Telecommunications Engineering
Universidad Militar Nueva Granada
Carrera 11 101-80, Bogota-Colombia
dora.ballesteros@unimilitar.edu.co, diego.renza@unimilitar.edu.co, u1400943@unimilitar.edu.co

Received June, 2015; revised December, 2015

Abstract. *In this paper we propose and validate a hypothesis that allows scrambling speech signals through an imitation process taking advantage of the Gaussian behavior of the speech signals. If a Gaussian noise signal has a similar entropy value and similar statistics of the speech signal, it is feasible that the speech signal imitates the behavior of the first one. At the output, the scrambled speech signal sounds like the Gaussian noise signal and its original content is destroyed. The length of the key (which relates the original and the final places of the samples of the speech signal) has the same length of the speech signal and it is generated in-situ within the imitation process. It is demonstrated that the proposed scheme works with unconditionally security because it satisfies the conditions of Shannon's perfect secrecy. Several tests were conducted in order to validate the hypothesis of imitation in terms of the residual intelligibility of the scrambled speech signal and the disorder grade of data.*
Keywords: Speech scrambling, Residual intelligibility, Covert communication, Audio steganography.

1. **Introduction.** Along with the massive use of internet and cellular, digital content has increased. Although many files are non-sensitive, others are classified by the user as confidential. In the latter, it is very important that only the intended recipient knows the content of information. One option for ensuring the confidentiality of information is covert communication. It means that the secret content is manipulated in two ways: destroying the original content by a scrambling process, or hiding the secret content in a host signal by a steganography process. In a hybrid solution, a hiding process is preceded by a scrambling process.

Speech scrambling is an old technique and this goes back some decades ago, at the beginning of the 20th century [1]. The aim is to disturb the parameters of the speech signal like amplitude, frequency, pitch or sequence. In this process if the elements affected are the samples sequence, the spectral components or the time-spectral components it is known as, Time Scrambling Permutation (TSP) [2, 3, 4], Frequency Domain Scrambling (FDS) [5, 6, 7, 8], or Time-Frequency Scrambling (TFS) [9, 10, 11], respectively. In these techniques, the new sequence is related to a key which typically is an input of the system. In the classical approach, the key is a pseudo-random sequence. Its main disadvantage is the lack of guarantee of non-repetitive sequences and therefore, the system is vulnerable. Hence, in the last years new proposals to obtain dynamic and non-predictive sequences

have been proposed. For example, chaotic maps are used in confusion and diffusion process [12, 13, 14] or cellular automata to generate the sequences [15]. However, most of scrambling schemes based on chaotic maps implies a high computational cost; and not all sequences obtained with cellular automata satisfy the condition of very low residual intelligibility of the scrambled speech signal. On the other hand, it is well known the Gaussian behavior of the speech signals [16, 17, 18]. It means that if two (or more) speech signals have different content (i.e. plaintext, intonation, gender of the speaker, among others) data distribution is similar between them. In the proposal of Ballesteros and Moreno, a scrambling scheme is presented taking advantage of the Gaussian behavior of the speech signals [19]. They propose a hypothesis of adaptation in which a speech signal with sensitive content can imitate a speech signal with non-sensitive content if and only if their histograms are similar [20, 21] . Then, the adapted speech signal (i.e. the scrambled speech signal) sounds like the non-sensitive speech signal, and the key (which is created in-situ) relates the original and final places of the samples. The main advantage of their proposal is the length of the key is equal to the length of the speech signal and all keys are likely. Therefore, it is ensured the unconditional security of the scheme. However, the disadvantage is that the non-sensitive speech signal (which is the target in the imitation process) must satisfy some conditions and it implies to have a big database of speech signals. Following the above concept, we propose a new hypothesis of imitation of speech signals that allows destroying the content of the original speech signal. In our case, the speech signal imitates a Gaussian noise speech signal which has similar statistics and entropy of the original one. With our proposal, we don't need a database of target signals because the Gaussian noise signals are created in-situ; but, the advantages of the proposal of [19] remains.

2. **Hypothesis Formulation and Statements.**

In this section we present a hypothesis of speech scrambling based on the entropy value.

Hypothesis: "If a speech signal with intelligible content has the same entropy of a Gaussian noise signal, the scrambled speech signal obtained by an imitation process would have unintelligible content".

The hypothesis is true under the following conditions:

1. Speech signals can be modeled as Gaussian functions.
2. It is feasible to obtain a Gaussian noise signal from desirable statistics.
3. If two signals have the same statistics within the same range, their PDF (Probability Density Function) are equal.
4. Scrambling process by imitation does not modify the PDF of the signal, neither its entropy nor statistics.

Proof. Entropy is an information content measurement of a signal. If the quantity of elements (i.e. samples in a digital speech signal) is finite, entropy can be calculated as follows:

$$H(X) = -\sum_{i} P(x_i)\ln(P(x_i)) \tag{1}$$

With $P(x_i)$ as the probability of occurrence of x_i. The unit of the above result is nat.

Let us consider the set of N amplitudes $[x_1\ x_2 \dots\ x_N] \in X$ and the set of probabilities $[p_{1_x}\ p_{2_x} \dots p_{N_x}] \in P$, satisfying:

$$P(x_i) = \ p_{1_x} \quad for \quad i = 1, 2, \dots N \tag{2}$$

Therefore,

$$H(X) = -\sum_{i} p_{i_x}\ln(p_{i_x}) \tag{3}$$

Let us choose two signals s_1 and s_2 with the same length, X and P. The difference between s_1 and s_2 lies on the place of their amplitudes. Denote,

$$p_{i_{s_1}} = p_{i_{s_2}} \quad for \quad i = 1, 2, \ldots N \tag{4}$$

Then we have,

$$H(S_1) = -\sum_i p_{i_{s_1}} \ln(p_{i_{s_1}}) = H(S_2) = -\sum_i p_{i_{s_2}} \ln(p_{i_{s_2}}) \tag{5}$$

Let us suppose s_1 is a speech signal and s_2 is a noisy signal, both with Gaussian behavior, same dynamic range and PDF. If follows that $H(s_1) = H(s_2)$.

With the scrambling process by imitation (it is explained in the next section), the scrambled speech signal, ss_1, is related to the Gaussian signal, as follows:

$$ss_1 = s_2 \tag{6}$$

Then, ss_1 is a noise Gaussian signal with the same content of s_2. If s_2 is an unintelligible signal, ss_1 is unintelligible signal, too. Further, we have:

$$H(ss_1) = H(s_2) \tag{7}$$

In a real case, the entropy of the Gaussian noise signal is very close to that of the speech signal. In a similar way, statistics and PDF of the Gaussian signal are similar to those of the speech signal, but not the same. Therefore, the scrambled speech signal will have similar behavior to the Gaussian noise signal, but with a slight difference.

Summarizing, the scrambled speech signal will have unintelligible content and a very close behavior of a Gaussian noise signal.

3. **The Proposed Scheme.** In this approach, the scrambled speech signal is obtained by an imitation process between the speech signal and the Gaussian noise signal. The most important requirement is the similarity in the behavior (entropy and statistics) of the speech signal and the Gaussian noise signal. In this section the scrambling and descrambling modules are explained together with a summary of the key generator characteristics and one example of the results.

3.1. **Scrambling module.** The input of this module is the speech signal which has intelligible content and sensitive information. At the output, the scrambled signal is obtained. This signal has unintelligible content and looks like a Gaussian noise signal. The steps of this module are: measurement, Gaussian noise generator and imitation process (Figure 1).

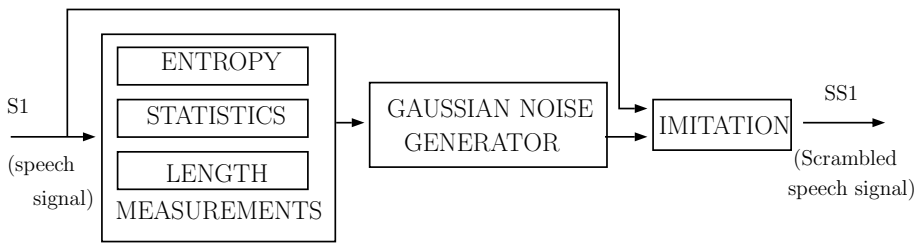


FIGURE 1. Block diagram of scrambling process.

3.1.1. *Measurement:* the aim of this block is to calculate the entropy of the speech signal, the four moments and the length. Entropy is obtained according to (1); the four moments are obtained, as follows:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N-1} \qquad \sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i-\mu)^2}{N-1}} \qquad sk = \frac{\sum_{i=1}^{N} (x_i-\mu)^3}{(N-1)\sigma^3} \qquad kt = \frac{\sum_{i=1}^{N} (x_i-\mu)^4}{(N-1)\sigma^4}$$

Where μ is the average of data, σ is the standard deviation, sk is the skewness and kt is the kurtosis.

3.1.2. *Gaussian noise generator:* with the values of the four moments of the speech signal, it is created a Gaussian noise signal with the same (or similar) entropy value of the speech signal. The output of this block is a signal with unintelligible content and with the same length of the speech signal.

3.1.3. *Imitation:* it takes advantage of the speech signal behavior. Since its histogram resembles Gaussian distribution [16, 17, 18], the speech signal can imitate a Gaussian noise signal. The process is developed, as follows:

a. The scrambled speech signal, ss_1, is initialized with zero values. Its length is the same as the secret message.
b. The secret key is initialized with zero values. Its length is the same as the secret message.
c. The highest values of the speech signal and the Gaussian noise signal are found. Positions are kept as I_{s_1} and I_{s_2}, respectively. The first value of ss_1 and the key are given as follows:

$$ss_1(I_{s_2}) = s_1(I_{s_1}) \tag{8}$$

$$key(I_{s_2}) = I_{s_1} \tag{9}$$

d. The second-highest values of both the speech signal and the Gaussian noise signal are found. Positions are kept as I_{s_1} and I_{s_2}, respectively. Equations (8) and (9) are carried out again.
e. The above procedure is repeated until all samples of the speech signal are relocated. At the output of this block, the scrambled speech signal sounds like a Gaussian noise signal.

3.2. **Descrambling module.** The aim of this module is to obtain the secret message from the scrambled speech signal. Since the scrambled speech signal is a relocated version of the secret message, its original content can be discovered only with the adequate key. If an unauthorized user intercepts the scrambled speech signal but he does not know the appropriate key, he is not able to obtain the secret content. The recovered secret message, rs_1, is obtained, as follows:

$$rs_1 = ss_1(key) \tag{10}$$

With the current proposal, the recovered secret message is equal to the original secret message if the scrambled speech signal is not manipulated in the transmission. But, with some level of manipulation, the content of the recovered secret message would be very close to the original secret message.

3.3. **Characteristics of the adaptive key generation.** With our proposal of key generation based on imitation process between the speech signal and the Gaussian noise signal, the following characteristics of the key are obtained:

i. Every pair of signals (secret message and Gaussian noise signal) has a unique key which related them.

ii. The key length is the same as the secret message length. Therefore, if the secret message has m samples, the length of the key is m.
iii. Data into the key are integer numbers in the range [1 m]. Every value between 1 to m is taken into account and none of them is repeated.
iv. If the length of the key is m, then there are up to $m!$ available keys. Furthermore, the higher the total number of samples, the higher is the number of available keys.
v. All keys are equally likely. It means none is more preferred than another.

3.4. **Example of the proposal.** With the purpose to illustrate the scrambling and descrambling modules in preliminary results, we select a speech signal with the plaintext "*today is a beautiful day*" with the following characteristics:
$H = 4.4509$, $\mu = -4.4179 \times 10^{-05}$, $\sigma = 0.1931$, $sk = 0.0607$, and $kt = 5.3695$.

With the above metrics, a Gaussian noise signal is generated with the following values:
$H = 4.2387$, $\mu = 1.1679 \times 10^{-04}$, $\sigma = 0.1913$, $sk = 0.1093$, and $kt = 5.1242$.

Then, the speech signal imitates the Gaussian noise signal with the procedure of section 3.1. At the descrambling module, with the scrambled speech signal and the key, the recovered speech signal is obtained. Figure 2 shows the speech signal, the Gaussian noise signal, the scrambled speech signal and the recovered speech signal. The similarity value between the speech signal and the adapted speech signal is 3.6956×10^{-05}, and between the target and the adapted is 0.9746. With the above results it is confirmed that the scrambled speech signal is highly similar to a Gaussian noise signal with illegible content. On the other hand, the imitation process is completely reversed and then, the recovered secret message is equal to the original speech signal (i.e. similarity equal to 1).

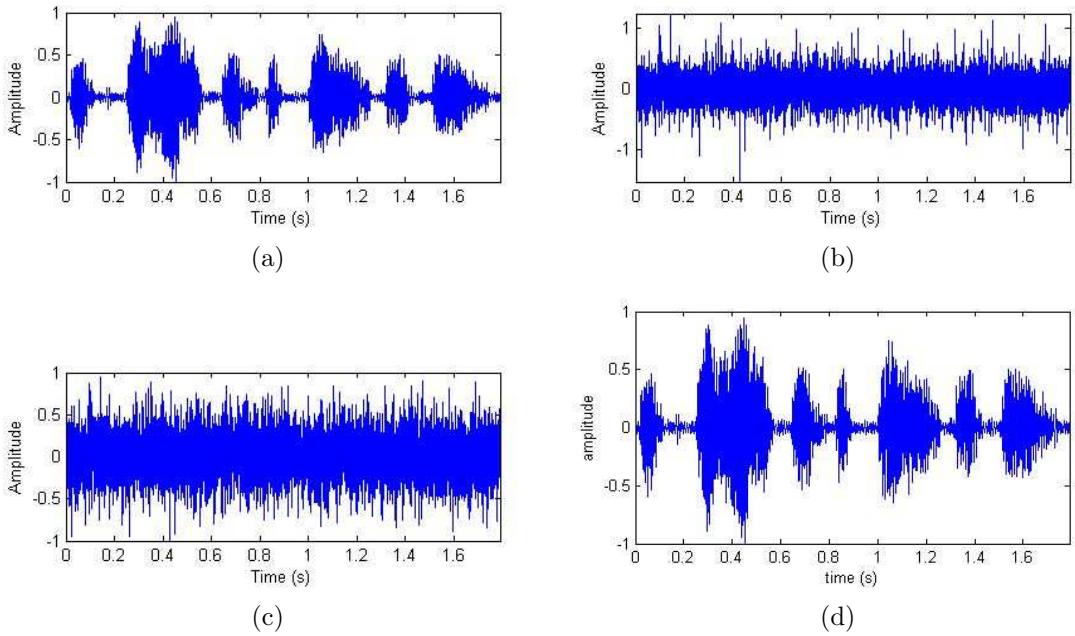


FIGURE 2. Example of the scheme: a) speech signal, b) Gaussian noise signal, c) scrambled speech signal, d) recovered speech signal.

4. **Results.** In this section we present the results of the residual in the scrambled speech signal, and we analyze the proposal in terms of perfect secrecy. For the experimental tests the following conditions were considered:

1. Speech signals are from five female and five male speakers.
2. There are five messages in English and five messages in Spanish language.
3. Every person speaks the messages in English and Spanish.
4. Ten trials were carried out by each imitation process.

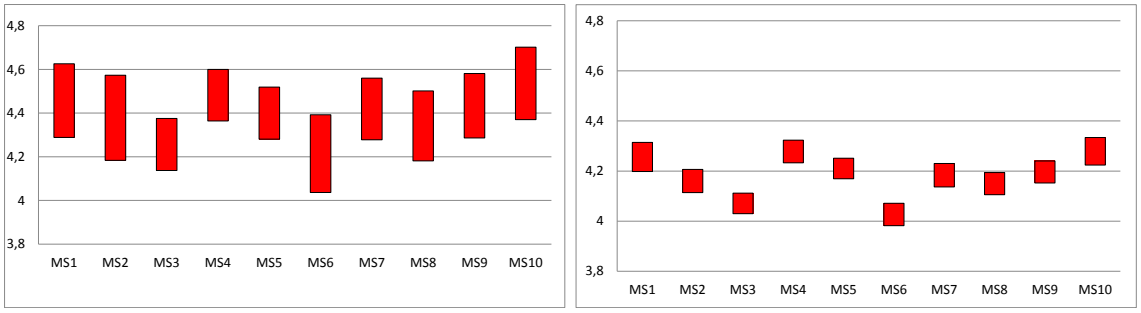
At the end, there are in total one thousand simulations. This is the result of multiply the total number of speakers (10) by the total number of speech messages (10) by the total number of trials by imitation process (10).

The steps carried out are:

a. The first step consists on eliminating the mute-time of every record. In total there are one hundred records.
b. Secondly, for the input record, entropy and statistics values are calculated.
c. Thirdly, a Gaussian noise signal is generated with the above values.
d. Trough imitation process, a scrambled speech signal is obtained. The original speech signal imitates the Gaussian noise signal. The *key* contains the relation of the original and final places of the samples.
e. To evaluate data repeatability, the third and fourth steps are run ten times.

4.1. **Behavior of the speech signal and the Gaussian Noise signal.** The first step in the proposed scrambling scheme is to obtain an adequate target signal to imitate. It means that the Gaussian noise signal should be the closest as possible to the speech signal, in terms of entropy and statistics.

Figures 3 and 4 show the results of entropy and kurtosis test of one thousand simulations. They are divided in ten cases which correspond to the ten messages of the test. Every case has one hundred simulations because there are ten speakers and ten trials of imitation by speaker. Cases are represented by the confidence range of 95%.



(a) (b)

FIGURE 3. Summary of entropy results: a) Speech signal, b) Gaussian noise signal.

According to Figure 3, the entropy of the Gaussian noise signal is similar to that of the speech signal. When entropy of the speech signal increases, entropy of the Gaussian noise signal increases too, and vice versa.

According to Figure 4, the gaussian noise signal follows the kurtosis of the speech signal. Therefore, if the kurtosis of the speech signal is high, the kurtosis of the Gaussian noise signal is high too, and vice versa.

Sumarizing, the gaussian generator provides a noise signal with similar behavior of the speech signal in terms of entropy and statistics, but with illegible content.

4.2. **Residual Intelligibility.** Residual intelligibility is the quantity of trace of the speech signal that remains in the scrambled speech signal. The purpose of any scrambling scheme is to provide very low residual intelligibility. Since by definition, the Gaussian noise

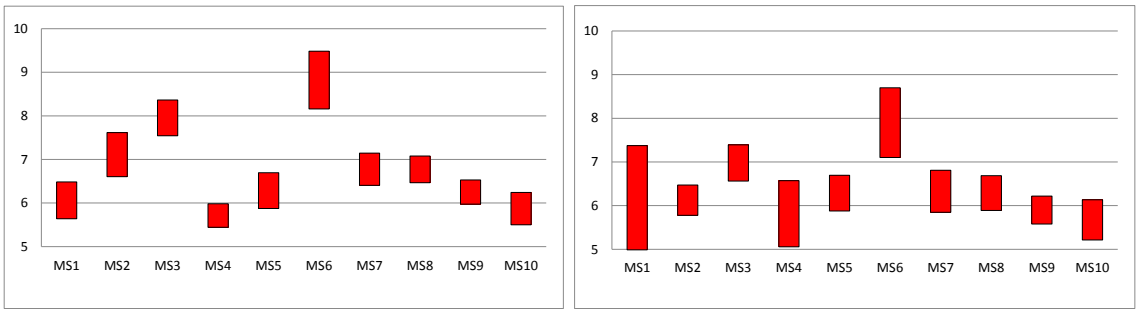


FIGURE 4. Summary of kurtosis results: a) Speech signal, b) Gaussian noise signal.

signal does not have intelligible content, the aim is satisfied if the scrambled speech signal is very close to the Gaussian noise signal. Therefore, the test is focused on confirming the similarity between these signals through two parameters: the SPCC (Squared Pearson Correlation Coefficient) and the GSD (Grade of Similarity of the Disorder). In the first, if the value is close to 1 it means that the Gaussian noise signal and the scrambled speech signal are highly correlated. In the second, if the value is close to 0 it means that the level of disorder is very similar between them, and therefore there is not intelligible content in the scrambled speech signal.

The squared correlation coefficient of two signals is a measure of their linear dependence. If each variable has N scalar observations, then the squared Pearson correlation coefficient is defined as,

$$SPCC = \frac{cov(s \times g)^2}{(\sigma_s)^2 \times (\sigma_g)^2} \tag{11}$$

Where cov is the covariance, $s(.)$ is the speech signal and $g(.)$ is the Gaussian noise signal.

The parameter GSD is calculated with the following equations:

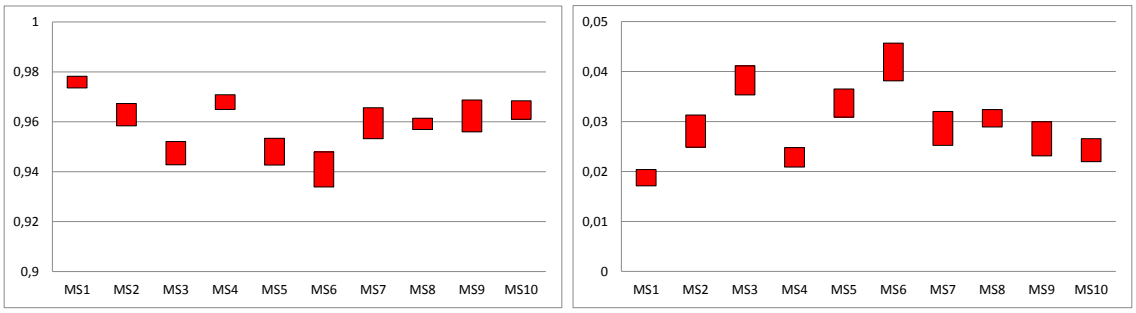
$$GSD = \frac{\overline{D_s} - \overline{D_g}}{\overline{D_s} + \overline{D_g}} \tag{12}$$

$$\overline{D_s} = \frac{\sum_{i=1}^{m-2} \sqrt{|s(i+1) - s(i)| + |s(i+1) - s(i+2)|}}{m-2} \tag{13}$$

$$\overline{D_g} = \frac{\sum_{i=1}^{m-2} \sqrt{|g(i+1) - g(i)| + |g(i+1) - g(i+2)|}}{m-2} \tag{14}$$

Where GSD is the similarity grade of disorder, m is the total number of samples of the signals, $\overline{D_s}$ is the average of disorder of the speech signal, and $\overline{D_g}$ is the average of disorder of the Gaussian noise signal.

Figure 5 shows the results of residual intelligibility in terms of $SPCC$ and GSD of one thousand simulations. According to Figure 5(a), the worst value of $SPCC$ is 0.93, which is enough to confirm the high similarity between the scrambled and the Gaussian noise signal. On the other hand, according to Figure 5(b), the level of disorder of the above signals is highly similar, and then the scrambled speech signal does not have residual content.



(a) (b)

FIGURE 5. Summary residual intelligibility test: a) SPCC, b) GSD.

4.3. **Perfect Secrecy.** In cryptography, the secret message is known as the plain-text, the scrambled message is known as cipher-text and the way to relate the plain-text with the cipher-text is the *key*. According to Shannon's theorem, a system that satisfies the concept of perfect secrecy is unconditionally secure, it means that no amount of computational power can help reveal the *key*, and therefore the secret content. In terms of perfectly key ambiguous, a system works with perfect secrecy if:

1. The sizes of the plain-text space (P), cipher-text space (C) and key space (K) are equal. It means if $|P| = |C| = |K|$.
2. All keys are equally likely
3. There is a unique key for every pair of plain-text and cipher-text.

The total number of Gaussian signals with the same value of entropy and statistics is equal to the total number of available permutations, it means $m!$, with m as the total number of samples. Since speech signals can be modeled as Gaussian signals, some of these signals have intelligible content (speech signal) and the others are Gaussian noise signals. The choice of an intelligible signal or noise signal from the set of signals is equally likely. Furthermore, through the imitation process, there is a unique *key* that relates the original speech signal and the Gaussian noise signal. With a wrong *key*, the recovered speech signal is part of the set of signals with intelligible content, but this signal does not correspond to the original speech signal. Summarizing, our proposal satisfies the three conditions of perfect secrecy and therefore the system is unconditionally secure.

4.4. **Comparison to related works.** In this section some of the most recent techniques of speech scrambling are compared in terms of the following characteristics: residual intelligibility in the scrambled speech signal, quality of the recovered secret message and security Table 1. The aim is to present the strengths and limitations of each method.

According to Table 1, there are some methods that allow perfect recovery of the secret message, but either they have low level of security or their residual intelligibility depends on external conditions. As opposed, our proposal has a good trade-off among residual intelligibility, quality of the recovered secret message and security. Comparing with our predecessor [19], we remain the strengths (i.e. security and quality of the recovered secret message) and overcome its weakness (i.e. residual intelligibility depends on the selected target speech signal) because the target speech signal is created in situ with zero trace of the original secret message.

5. **Conclusion.** In this paper a speech scrambling scheme has been proposed, which uses an imitation process between the speech signal (with sensitive content) and a Gaussian

TABLE 1. Comparative results

Work	**Method**	**Residual intelligibility**	**Quality of the recovered secret message**	**Security**
[22]	High dimension transformation	It is very low but depends on the selected sequence	The process is completely reversed	It does not resist brute force attacks because the block size is short
[15, 23]	Chaotic sequences	It is very low	The recovered speech signal is very similar to the original one. It needs synchronization	Crypto systems based on chaos can be broken [23]
[24, 25]	Cellular automata	It depends strongly on initial conditions (i.e. transition rules, neighborhood rules, number of generations)	The process is completely reversed	The system can be broken. In [26] the key space is fixed and short (i.e. 2^{96}) and does not depend on the speech length
[19]	Imitation of non-sensitive speech signal	It can be very low but depends on the selected target speech signal. It needs a huge database of target speech signals	The process is completely reversed	It works with perfect secrecy
Our proposal	Imitation of Gaussian Noise signal	It is zero trace of the original content. The results are not-dependent of the secret message	The process is completely reversed	It works with perfect secrecy

noise signal. It is feasible because the speech signal can be modeled as a Gaussian function. The characteristics of our proposal are:

i. A Gaussian noise signal with similar entropy and statistics values of the speech signal is used as the target signal in the imitation process.
ii. The speech signal imitates the Gaussian noise signal and then the intelligible content is destroyed in the process. The scrambled speech signal sounds like a Gaussian noise signal with zero residual intelligibility.
iii. The *key* is obtained in-situ.
iv. The Gaussian noise signal is obtained in-situ. It means there is not necessary a database with keys or target speech signals.
v. The *key* length is equal to the total number of samples of the speech signal. There is a unique key by every pair of signals (speech signal and scrambled speech signal) and all keys are equally likely. Therefore, it is established that the system works with perfect secrecy, in a similar way of the scheme proposed in [19].

In relation to our predecesor:

A. Our proposal does not need an external database of target speech signals.
B. Our residual intelligibility is zero in all cases. It does not depend on the selected target speech signal.
C. The strengths remain (high security, quality of the recovered secret message).

In sumary, our proposal has a very good trade-off among security, residual intelligibility and quality of the recovered secret message.

6. **Acknowledgment.** This work is supported by the "Universidad Militar Nueva Granada-Vicerrectoría de Investigaciones" under the grants INV-ING-1910 and INV-ING-1768 of 2015.

REFERENCES

[1] N. MacKinnon, "The development of speech encipherment," *Radio and Electronic Engineer*, vol. 50, no. 4, pp. 147–155, 1980.
[2] S. Kak, "Encryption of signals using data transpositions," in *Proceedings of the Institution of Electrical Engineers*, vol. 125, pp. 1327–1328, IET, 1978.

[3] E. Mosa, N. W. Messiha, and O. Zahran, "Random encryption of speech signal," in *Computer Engineering & Systems, ICCES. International Conference on*, pp. 306–311, IEEE, 2009.
[4] V. Phillips, M. Lee, and J. Thomas, "Speech scrambling by the re-ordering of amplitude samples," *Radio and Electronic Engineer*, vol. 41, no. 3, p. 99, 1971.
[5] E. Del Re, R. Fantacci, and D. Maffucci, "A new speech signal scrambling method for secure communications: theory, implementation, and security evaluation," *Selected Areas in Communications, IEEE Journal on*, vol. 7, no. 4, pp. 474–480, 1989.
[6] Y. Lim, J. Lee, and S. Foo, "Quality analog scramblers using frequency-response masking filter banks," *Circuits, Systems and Signal Processing*, vol. 29, no. 1, pp. 135–154, 2010.
[7] A. Matsunaga, K. Koga, and M. Ohkawa, "An analog speech scrambling system using the fft technique with high-level security," *Selected Areas in Communications, IEEE Journal on*, vol. 7, no. 4, pp. 540–547, 1989.
[8] D. Tseng and J. Chiu, "An ofdm speech scrambler without residual intelligibility," in *TENCON IEEE Region 10 Conference*, pp. 1–4, IEEE, 2007.
[9] J. F. de Andrade, M. L. R. de Campos, and J. A. Apolinario, "Speech privacy for modern mobile communication systems," in *Acoustics, Speech and Signal Processing, ICASSP. IEEE International Conference on*, pp. 1777–1780, IEEE, 2008.
[10] F. Ma, J. Cheng, and Y. Wang, "Wavelet transform-based analogue speech scrambling scheme," *Electronics Letters*, vol. 32, no. 8, pp. 719–721, 1996.
[11] S. Sadkhan, N. Abdulmuhsen, and N. F. Al-Tahan, "A proposed analog speech scrambler based on parallel structure of wavelet transforms," in *National Radio Science Conference*, pp. 1–12, IEEE, 2007.
[12] R. Gnanajeyaraman, K. Prasadh, *et al.*, "Audio encryption using higher dimensional chaotic map," 2009.
[13] M. Ashtiyani, P. M. Birgani, and S. K. Madahi, "Speech signal encryption using chaotic symmetric cryptography," *J. Basic. Appl. Sci. Res*, vol. 2, no. 2, pp. 1678–1684, 2012.
[14] S. M. Alwahbani and E. B. Bashier, "Speech scrambling based on chaotic maps and one time pad," in *Computing, Electrical and Electronics Engineering (ICCEEE), International Conference on*, pp. 128–133, IEEE, 2013.
[15] A. Madain, A. L. A. Dalhoum, H. Hiary, A. Ortega, and M. Alfonseca, "Audio scrambling technique based on cellular automata," *Multimedia Tools and Applications*, vol. 71, no. 3, pp. 1803–1822, 2014.
[16] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, no. 6, pp. 1109–1121, 1984.
[17] I. Cohen, "Modeling speech signals in the time–frequency domain using garch," *Signal Processing*, vol. 84, no. 12, pp. 2453–2459, 2004.
[18] J. W. Shin, J.-H. Chang, and N. S. Kim, "Statistical modeling of speech signals based on generalized gamma distribution," *Signal Processing Letters, IEEE*, vol. 12, no. 3, pp. 258–261, 2005.
[19] D. M. Ballesteros L and J. M. Moreno A, "Speech scrambling based on imitation of a target speech signal with non-confidential content," *Circuits, Systems, and Signal Processing*, pp. 1–24, 2014.
[20] D. M. Ballesteros L and J. M. Moreno A, "On the ability of adaptation of speech signals and data hiding," *Expert Systems with Applications*, vol. 39, no. 16, pp. 12574–12579, 2012.
[21] D. M. Ballesteros L and J. M. Moreno A, "A bit more on the ability of adaptation of speech signals," *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 66, pp. 82–90, 2013.
[22] H. Li, Z. Qin, L. Shao, S. Zhang, and B. Wang, "Variable dimension space audio scrambling algorithm against mp3 compression," in *Algorithms and Architectures for Parallel Processing*, pp. 866–876, Springer, 2009.
[23] S. N. George, N. Augustine, and D. P. Pattathil, "Audio security through compressive sampling and cellular automata," *Multimedia Tools and Applications*, pp. 1–25, 2014.
[24] L. J. Sheu, "A speech encryption using fractional chaotic systems," *Nonlinear Dynamics*, vol. 65, no. 1-2, pp. 103–108, 2011.
[25] Z.-P. Jiang, "A note on chaotic secure communication systems," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 49, no. 1, pp. 92–96, 2002.
[26] A. B. Orue, V. Fernandez, G. Alvarez, G. Pastor, M. Romera, S. Li, and F. Montoya, "Determination of the parameters for a lorenz system and application to break the security of two-channel chaotic cryptosystems," *Physics Letters A*, vol. 372, no. 34, pp. 5588–5592, 2008.