Secret Keys Distillation using Speech Signals and Discussion over Public Authenticated Channel

Jelica Radomirović, Milan Milosavljević, and Aleksandra Krstić

***Abstract*— This paper discusses a system for generating and distributing secret cryptographic keys based on the principle of common randomness and discussion over an authenticated public channel. The use of speech as a source of common randomness is one possibility and to our knowledge the first for this type of system. We consider different reconciliation algorithms and compare them with experiments. Experimental results show that it is possible to generate information-theoretical secret keys with rates between 3 and 5%. This result proves the practical feasibility of absolutely secret autonomous cipher systems with speech control.**

***Index Terms*—secret key; distillation; symmetric cryptography; speech signal; reconciliation; privacy amplification; secret key rate;**

I. Introduction

The information-theoretical approach to the analysis and synthesis of cipher systems came into focus with the availability of quantum computers in the near future. The classical result of this approach states that the entropy of secret keys in a cryptographic system must be no less than the entropy of plaintext [1]. As is well known, systems designed in this way are resistant to the unlimited computing resources of the adversary, and thus to cryptanalysis based on quantum computers [2].

From the point of view of generating and distributing high-quality secret keys, special attention is drawn to the fundamental results of Alswede and Csiszar [3], Maurer [4], and Csiszar and Narayan [5]. The basic idea of information-theoretical approach in these results is to identify and use mutually correlated signals available to legitimate parties.

Depending on the location of the source of common randomness, there are two approaches, [3]:

(i) Secret key extraction from sources independent of communication channels (source model),

(ii) Secret key extraction from existing communication channels (channel model).

The difference between these two models is how the parties observe the initial sequence. While in the source model, random source is controlled by nature, in the channel model, one of the parties governs the input of a noisy channel (independent of the main channel) while others observe the output.

In this paper, we will analyze the possibility of extracting cryptographic keys from a speech signal, applying an approach based on the source model.

In Section 2, the basic blocks of the proposed secret key generation system will be presented, in two variants: (i) when the input is a speech signal and (ii) when the input is a residual speech signal, filtered by an adaptive linear predictive model [6].

In Section 3, the information and statistical characteristics of this source will be analyzed and the key parameters of the sequential procedure for extracting secret keys will be identified, separately for each of the phases: Advantage Distillation (AD), Information Reconciliation (IR) and Privacy Amplification. -PA).

In Section 4 we present the results of the experiment of obtaining secret keys for all pairs (Alice, Bob) of legitimate participants for 5 speakers, one of which was chosen as an eavesdropper (Eve).

The Conclusion discusses the upper limits of the rate of generating secret keys and the possibility of further improving the performance of the proposed system.

II. Discrete Memoryless Source

As illustrated in Figure 1, a source model for secret-key agreement represents a situation in which three parties, Alice, Bob, and Eve, observe the realizations of a DMS - Discrete Memoryless Source (XYZ, P_{XYZ}) with three components.

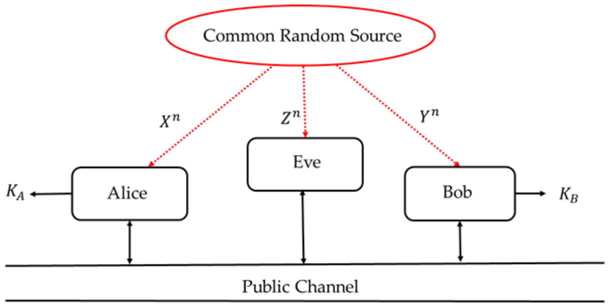

Fig. 1. Secret-key Agreement by Public Discussion from Common Randomness [4].

Jelica Radomirović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia and with Vlatacom Institute of High Technologies, 5 Bulevar Milutina Milankovića, 11070 Belgrade, Serbia (e-mail: jelica.radomirovic@vlatacom.com).

Milan Milosavljević is with Singidunum University, 32 Danijelova, 11000 Belgrade, Serbia (e-mail: mmilosavljevic@singidunum.ac.rs).

Aleksandra Krstić is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: amarjanovic@etf.bg.ac.rs).

The DMS is assumed to be outside the control of all parties, but its statistics are known. By convention, component X is observed by Alice, component Y by Bob, and component Z by Eve. Alice and Bob's objective is to process their observations and agree on a key K about which Eve should have no information.

Alice and Bob can exchange messages over a noiseless, two-way, public and authenticated channel. That is, all messages are overheard by Eve and the existence of the public channel does not provide Alice and Bob with an explicit advantage over Eve. The rules by which Alice and Bob compute the messages they exchange over the public channel and agree on a key define a four-stage key distillation strategy, [4]:

1. Randomness sharing. Alice, Bob, and Eve observe *n* realizations of a DMS (XYZ, P_{XYZ}).
2. Advantage distillation. If needed, Alice and Bob exchange messages over the public channel to process their observations and to "distill" observations for which they have an advantage over Eve.
3. Information reconciliation. Alice and Bob exchange messages over the public channel to process their observations and agree on a common bit sequence.
4. Privacy amplification. Alice and Bob publicly agree on a deterministic function they apply to their common sequence to generate a secret key.

The largest achievable key rate is defined as the key capacity and is given by

$$C_K = max\{I(X;Y), I(X:Y|Z)\}, \qquad (1)$$

where I(X;Y) denotes mutual information between X and Y, while I(X:Y|Z) denotes the same quantity conditioned by Z. In the special case, when Eva is totally independent of Alice and Bob, or equivalently, when Z is independent of X and Y, maximal key capacity is equal to

$$C_{K\,max} = I(X;Y). \qquad (2)$$

In this work, we use speech the signals of participants as DMS of the proposed system, see Fig.2.

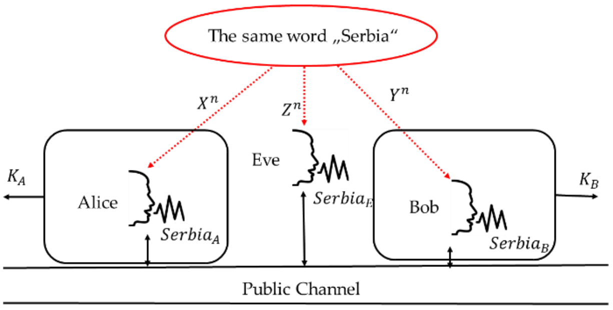

Fig. 2. Secret-key Agreement by Public Discussion based on the speech signals obtained by pronouncing the word "Serbia".

III. System Architecture

As already mentioned, we will analyze two DMS, the first one corresponding to the original speech signal, and the second one corresponding to the residual signal. Residual DMS is obtained after inverse filtering by an adaptive linear autoregressive model, estimated every 10 ms of input speech, see Fig.3.

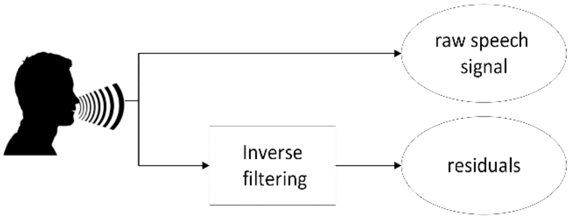

Fig. 3. Two different DMS based on the same input speech signal

The general architecture of the system is given in Fig.4. Speech input (or residual) is transformed into binary DMS by a non-uniform quantization, based on estimating the probability density function of input samples.

Advantage Distillation (AD) blocks are used to eliminate the advantage that the eavesdropper may have over legitimate parties. In that case, Eve knows more information about Alice's initial bit string, than Bob does. We will use the bit pair advantage distillation/degeneration protocol [7]. Algorithms are based on the exchange of parity information of 2-bit blocks and the elimination of one bit of each block to ensure security. After this step, the number of different bits of the sequence is reduced. The number of non-matching bits of the eavesdropper sequence decreases but much slower than for the legitimate parties. The protocol runs for several rounds until the sequences differ only in a few bits.

The Information Reconciliation (IR) is intended to correct the remaining erroneous bits in the Alice and Bob sequences. The most popular IR protocols are Winnow [7] and Cascade [8] protocols. They are based on the exchange of block parity information until the error of a certain block is detected. For each block parity query, some bits are deleted to ensure security. In the end, we get the same sequence on legitimate sides that represents the secret key. Even though we tried to ensure privacy by deleting potentially compromised bits, the eavesdropper has still gained some information about the secret key. To make the secret key absolutely secure, we proceed to the next step of our system, privacy amplification (PA).

In Privacy Amplification (PA) block sequences are transformed such that m bits are discarded due to the eavesdropper's knowledge. One of the possible transformations is a hash function $g: \{0,1\}^n \rightarrow \{0,1\}^r$ where n is the length of a sequence before PA and r is the length after PA, i.e., the length of the final secret key. It is common to use so-called universal hash functions, such as random *r x n* matrices, over GF(2), [9].

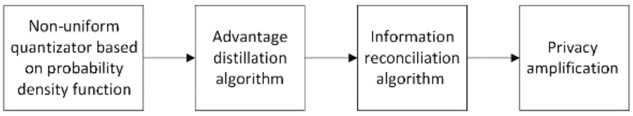
Non-uniform quantizator based on probability density function
Advantage distillation algorithm
Information reconciliation algorithm
Privacy amplification

Fig. 4. Secret key agreement algorithm [3]

IV. Results

Raw speech signal and its residuals, obtained from inverse filtration of AR model of 10th order, were used for experimental proof of the proposed system, Fig 5. Four participants recorded the word ‘Srbija’ for two seconds. Beginning. as well as the end of the word, were determined so the initial signal is reduced to 0.6 second length. The recordings were sampled at 44.1 kHz.

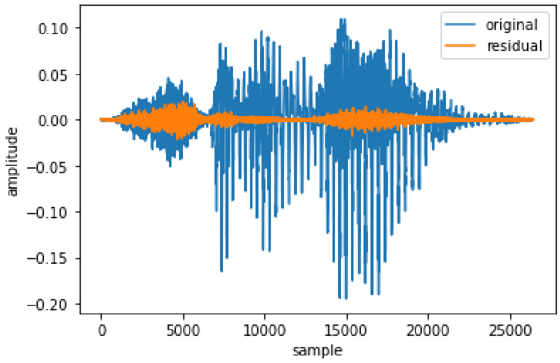

Fig. 5. Comparison of source signals

Normalized Hamming distance is an appropriate metric for measuring the difference between two binary sequences. Depending on the number of bits we use to quantize continuous signal we get more or less similar sequences. That directly affects how much information circulates over a public channel and how long is the final secret key length. The key rate is an indicator of how much of the sequence at the beginning is useful

$$\text{key rate}=\frac{\text{final length of secret key}}{\text{length of input sequence}}*100\ [\%].$$

In Fig 6. the key rate in function of normalized Hamming distance is presented. In order to determine the optimal value for the number of quantization bits, we try five different values 6,8,10,12, and 14.

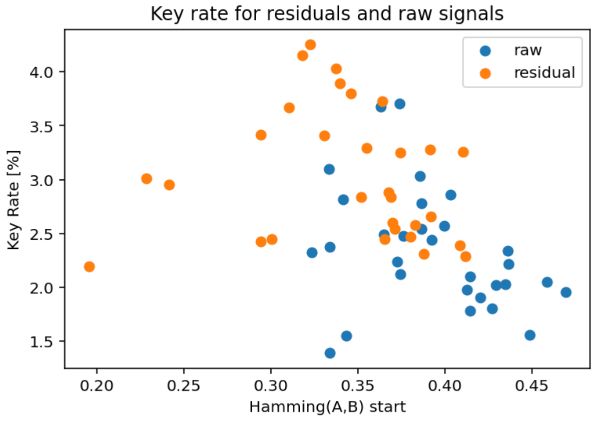

Fig. 6. Comparison of key rate for source signals. A and B denote legitimate parties

From Fig.6 can be seen that residuals are closer to each other than raw signals because of the corresponding higher key rate.

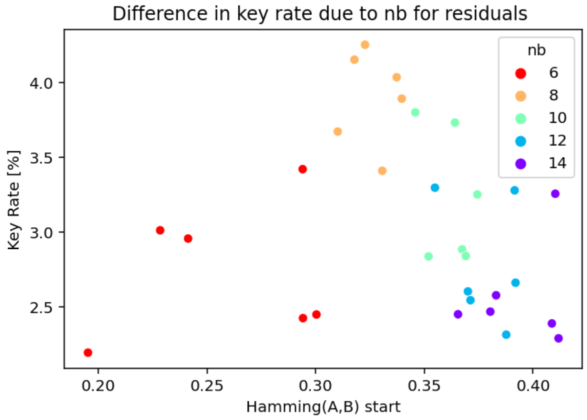

Fig. 7. How the number of quantization bits affects distance at the beginning

As shown in Fig 7., the highest rate is achieved for normalized Hamming distance between 0.3 and 0.35, which corresponds to 8 quantization bits.

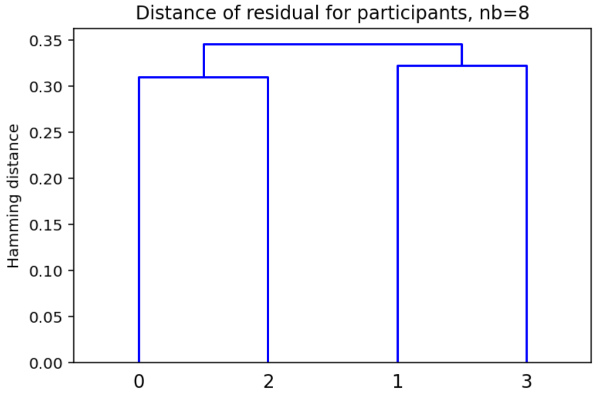

Fig. 8. Dendrogram represented distance of all sequences

In Fig 8. we use a dendrogram to show how close to each other are the participant sequences. The dendrogram was obtained as a result of hierarchical cluster analysis by the Ward method [10]. To compare Cascade and Winnow algorithms we conduct 4 experiments, 2 for each, that is one for residual signals and one for speech signals. Results are represented in table 1.

TABLE I
Experimental Results

		Final key size	Key rate [%]
Winnow	Residual	10116.83 ± 788.93	4.78 ± 0.37
	Raw	8670.30 ± 1033.87	4.10 ± 0.49
Cascade	Residual	8432.24 ± 624.18	3.98 ± 0.29
	Raw	6833.72 ± 1433.26	3.22 ± 0.68

Based on the results we conclude that the Winnow algorithm is a better choice for reconciliation because it gives an almost 1% higher key rate. If we compare two DMS, the one corresponding to the residual signal gives longer secret keys due to a smaller normalized Hamming distance at the beginning, for the same pair of sequences. Final hamming for all sequences and all experiments toward the eavesdropper after PA is ~0.5. In other words, all the information that leaked through the public channel will not reveal anything to the eavesdropper about the distilled secret key.

V. Conclusion

In this work we proposed a speech based secret key agreement system with message transmission over a public channel. The proposed system can distill secret keys from speech signals, with the key rate of up to 5%, and with negligible information leakage to an eavesdropper. This opens up the possibility of practical realization of absolutely secret cipher systems controlled by voice. Such systems can be used both in the security services for critical information and communication infrastructure of the government, as well as in commercial applications.

Future work will include generalization in terms of the largest achievable key rate and a testing of the proposed system on more participants as well as more different spoken words.

VI. Acknowledgment

The authors would like to thank the Vlatacom Institute of High Technologies, where the research was done as part of the project Prj_164.

References

[1] Shannon C.E., “Communication theory of secrecy systems”. *BSTJ*, vol. 28, no. *4*, pp. *656–715*. October 1949.

[2] Wolf S., “Unconditional Security in Cryptography”, in *Lectures on Data Security: Modern Cryptology in Theory and Practice, Lecture Notes in Computer Science*, Berlin, 1999, vol. 1561, pp. *217–250*.

[3] Ahlswede R., Csiszar I., “Common randomness in information theory and cryptography, Part I: Secret sharing”, *IEEE Transaction on Information Theory*, vol. 39, pp. *1121–1132*, 1993.

[4] Maurer U., “Secret Key Agreement by Public Discussion from Common Information”, *IEEE Transaction on Information Theory*, vol. 39, no. *3*, May, 1993.

[5] Csiszar I., Narayan P., “Secrecy capacities for multiple terminals”, *IEEE Transaction on Information Theory*, vol. 50, pp. *3047–3061*, 2004.

[6] Kovačević B., Milosavljević M., Veinović M., “Robust Digital Processing of Speech Signals“, Springer, 2017.

[7] Wang Q., Wang X., Lv Q., Ye X., Luo Y., You L., “Analysis of the information theoretically secret key agreement by public discussion”, *Security and Communication Networks*, vol. 8, January, 2015.

[8] Reis A., “Quantum Key Distribution Post Processing - A Study on the Information Reconciliation Cascade Protocol”. Master’s Thesis, Faculdade de Engenharia, Universidade do Porto, Porto, Portugal, 2019.

[9] Bennett C.H., Brassard G., Crepeau C., Maurer U. “Generalized privacy amplification”. *IEEE Transaction on Information Theory*, vol. 41, pp. 1915–1923, 1995.

[10] sklearn.cluster.Ward – scikit-learn 0.15-git documentation. https://scikit-learn.org/0.15/modules/generated/sklearn.cluster.Ward.html