4th International Conference on Advanced Technologies
For Signal and Image Processing – ATSIP' 2018
March 21-24, 2018 – Sousse, Tunisia

Cancelable Speaker Verification System Based on Binary Gaussian Mixtures

Aymen Mtibaa
Télécom SudParis, SAMOVAR CNRS
Paris-Saclay University
ENIS, ATMS
Sfax University
aymen.mtibaa@telecom-sudparis.eu

Dijana Petrovska-Delacrétaz
Télécom SudParis, SAMOVAR CNRS
Paris-Saclay University
Evry, France
dijana.petrovska@telecom-sudparis.eu

Ahmed Ben Hamida
ENIS, ATMS
Sfax University
Sfax, Tunisia
ahmed.benhamidag@gmail.com

***Abstract*—Biometric systems suffer from non-revocabilty. In this paper, we propose a cancelable speaker verification system based on classical Gaussian Mixture Models (GMM) methodology enriched with the desired characteristics of revocability and privacy. The GMM model is transformed into a binary vector that is used by a shuffling scheme to generate a cancelable template and to guarantee the cancelabilty of the overall system. Leveraging the shuffling scheme, the speaker model can be revoked and another model can be reissued. Our proposed method enables the generation of multiple cancelable speaker templates from the same biometric modality that cannot be linked to the same user. The proposed system is evaluated on the RSR2015 databases. It outperforms the basic GMM system and experimentations show significant improvement in the speaker verification performance that achieves an Equal Error Rate (ERR) of 0.01%.**

***Index Terms*—speaker verification, cancelable speaker system, revocability, privacy**

I. Introduction

Biometric applications are gaining more and more popularity. Regarding the progress of biometric technologies, voice biometrics [1] has the advantage that it can be applied in diverse situations that cover mobile commerce and transactions due to the ubiquity of the microphone. However, there are some issues related to biometric systems [2]. In available biometric systems, the generated biometric templates are similar in different applications because of the employment of the same biometric feature and the same computational procedure. In speaker recognition systems, the speaker model formed from the same voice for different applications is similar. If the speaker model of an application is stolen, it can be exploited to access other applications. This is presenting a threat to privacy. In addition, since biometric characteristics are eternally associated with the user, it is impossible to replace the compromised template with a new one which means non-revocability [3]. Therefore, protection of biometric data is considered as an important requirement to avert privacy and security threats [4]. Currently, biometric templates are protected by storing them in an encrypted place. However, during the biometric comparison, they need to be decrypted inducing a weakness.

Since 1990, new methodologies have emerged to protect the biometric template in a better way. Cancelable biometric systems [5] are one of the methods that are used to protect the biometric template. Cancelable biometric systems transform the original biometric data with user specific transformations in such a way that it is difficult to recover the original biometric data. Contrary to template protection by encryption algorithms, comparison of the cancelable template is performed in the transformed domain. A perfect template protection scheme should satisfy the following requirements [6]:

Revocability: The protected biometric template should be able to be revoked and renewed to replace the compromised template. In addition, the new protected template should not match with the old template so that the system can reject a user who provides the old template.

Performance: The biometric template protection system should not degrade the accuracy of the underlying baseline biometric system.

Non-invertibility: It should be computationally infeasible to recover the original biometric template from the protected template.

Unlinkability: Given the same biometric data, it must be feasible to generate different versions of protected biometric templates in a way that they cannot be linked to each other or to the subject from which they were derived.

In accordance with the ISO/IEC 24745 on biometric information protection [7], the architectural aspects of cancelable biometric systems are presented in Figure 1. In the enrollment phase, a biometric feature is extracted. Then, it is taken by a pseudonymous identifier encoder (PIE) as an input to generate a pseudonymous identifier (PI) and corresponding auxiliary data (AD). PI represents the protected identity of biometric features and AD represents a subject specific data. Both PI and AD constitute the protected template. Recall that PI and AD are stored and can be separated from each other. After enrollment, the biometric feature data is deleted. At verification phase, the pseudonymous identifier recorder (PIR) extracts the biometric features test and takes the auxiliary data of the user as input and generates a pseudonymous test identifier (PI*). At last, a pseudonymous identifier comparator (PIC) compares the PI generated during enrollment and PI* and returns a similarity score.

Despite the huge research that has been done to protect biometric template, most of the used templates are in terms

978-1-5386-5239-8/18/$31.00 ©2018 IEEE

of iris, fingerprint and face modalities. In this paper, we are interested in providing cancelable (also denoted as revocable) speaker verification system.

The structure of this paper is as follows: section II reviews related works on speaker recognition based on biometric template protection. The proposed binary UBM-GMM-based cancelable speaker verification system is described in section III. Experimental protocols and results are presented in section IV. Conclusion and perspectives are given in section V.

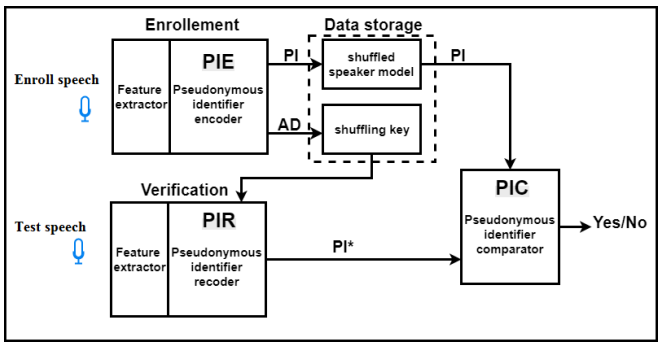

Fig. 1. Architecture for cancelable biometric systems

II. Related work

Most of the research on cancelable biometrics is proposed for iris, fingerprint and face modalities. Ratha et al. [8] made the first essay to protect the privacy of user with the transformation of biometric features. Recently, authors in [9] reported three proposal transformations: i) cartesian, ii) polar and iii) functional. All of them are realized on fingerprints. Their analysis shows that this transformation resists to different attacks in order to recover the original template. However, the performance accuracy degrades in the transformed domain which represents a strong limitation of their proposal in practical applications. Further, we enumerate the proposed cancelable solutions for speaker verification applications.

Teoh et al. [10] propose a cancelable voice biometric system based on probabilistic random projection. A two dimensional analysis for the principal component is applied to the feature matrix. Then, the feature matrix is projected in a random projection process. The projected matrix was then fed into a Gaussian Mixture Model (GMM). Their proposed scheme protection has shown good performance accuracy especially in the stolen key scenario. However, assuring unlinkability and revocability is still unclear in their work.

Based on the fuzzy vault scheme [11], Xu et al. [12] proposed a cancellable voice template protection. Chaff points are added to the Mel Frequency Cepstral Coefficient-MFCC matrix to compose a vault and to make hard for an adversary the separation of genuine points from the MFCC. While the authentication the server has difficulty with the chaff points separation and fails to compare protected template in transforming domain. For this, a prime accumulator is used to separate the genuine points from the chaff points. As drawbacks to this approach, after using prime accumulator, the server obtains the features MFCC in clear. Then, if an adversary compromises the prime accumulator, it will be easy to determine genuine points (MFCC). Moreover, the biometric evaluation metrics for cancelable speaker verification are not taken into account.

The difficulty of voice biometrics is that speaker recognition is based on statistical models that represent the user. This is not the case for iris, face or fingerprints, where features can be directly extracted and the user is represented by the so called template. Recent work presents binarization methods of speaker models to integrate protection scheme.

Paulini et al. [13] introduce binarization template called Multi bit Allocation to treat this issue. They proposed methods based on GMM-UBM speaker recognition system to extract discriminative binary feature applicable for template protection. Performance analysis for this binary feature shows slight degradation compared to the GMM-UBM baseline.

Bileb et al. [14] proposed another protection scheme for speaker recognition based on the binarization approach mentioned in [13]. The proposed scheme binarizes the supervector derived from the universal background models and a fuzzy commitment scheme is used as a basis for the template protection. However, the drawback of this system is that the performance degrades compared to the baseline system, the equal error rate increases from 3.4% to 5.42%.

We can conclude that in the above cited research the biometric performance of the cancelable system degrades compared to the baseline speaker verification system. Also, they don't take into consideration the complete set of evaluation metrics that need to be validated for a cancelable biometric system. Therefore, our approach is to propose a cancelable speaker verification system that improves the performance and includes the mentioned missing metric evaluations.

III. Cancelable Speaker Verification System Based on Binary UBM-GMM Model

In recent research, different methods are proposed to present voice biometrics models in binary format [15] [16]. In this paper, we adapt the binarization scheme from [17]. In fact, while a binary Key Background Model is used to address the acoustic space in their work, we leverage the GMM-UBM [18] based speaker verification system to serve as a baseline to obtain a binary representation of voice features. Given a large set of speech data gathered from hundreds of speakers, an universal background model UBM is trained with the Expectation-Maximization (EM) algorithm. During the enrollment, a specific speaker GMM model is derived from the UBM by adapting the UBM mean to the speaker's training utterances using the MAP (Maximum A Posteriori) adaptation. Then, we specify a binary vector of N bits length as speaker binary presentation, where N is the number of Gaussian mixtures in the GMM model. Each bit in the binary vector is linked with a Gaussian Mixture λ. We also define an accumulator vector of the same length as the binary vector initialized to 0. Given the speaker's utterance, for each acoustic frame in the feature vector $X[n]$, we compute its

likelihood $lkld$ given each of the Gaussians λ in the GMM model and select the top $\theta 1$ percent Gaussians with highest likelihood values. For the selected Gaussians, we increase by 1 the corresponding accumulator vector positions. Last step is the conversion of the accumulator vector to a speaker binary vector by setting the top $\theta 2$ percent positions in the accumulator vector with highest values to 1, and to 0 when otherwise [17].

After the binarization step, the speaker binary vector is transformed using a specified shuffling key for each user to generate the protected template. The shuffling scheme was initially proposed by [19]. In this work, we apply this shuffling scheme for speaker verification to induce revocability. The introduced shuffling scheme necessitates a binary shuffling key K of L bit length. Therefore, it is a two factor authentication, speech and shuffling key. The speaker binary vector is segmented into L blocks of equivalent lengths. To start the shuffling, these L blocks of binary vector are lined up with the L bits of the shuffling key K. In the next stage, two distinct parts containing binary vector features are created. The first part includes all blocks corresponding to positions where the value of shuffling key bit is one. The rest of blocks are all taken in the second part. These two parts are concatenated to form the shuffled binary data which is considered as the protected template. Figure 2 illustrates the steps to create the cancelable speaker verification system. The efficiency of this scheme is shown by its ability to affect only the alignment not the values of the binary vector. This is an important point because each value in the speaker binary vector is the projection of the acoustic location of each acoustic frame from the feature space into the space of GMM Gaussians.

To compare between two shuffled binary vectors a and b, a dissimilarity score s (Eq.1) is obtained by computing their Hamming distance as follows:

$$s(a,b) = 1 - \frac{\sum_{i=1}^{N} (a[i] \wedge b[i])}{N} \quad (1)$$

Where $\wedge$ is the operator of AND logic between any two bits.

IV. Experimental Protocols and Results

In this section, databases along with the experimental steps and results are described. The TIMIT database [20] is used to tune the system and to study the influence of $\theta 1$ and $\theta 2$ parameters for the binary representation of the speech data. Part1 of RSR2015 database [21] is used to evaluate the proposed system. A comparative study is performed with the baseline classical UBM/GMM system.

A. Databases and Protocols

TIMIT database contains a total of 6300 sentences, ten sentences spoken by each of the 630 speakers (438 males and 192 females) of eight major dialects of American English. The textual material in the TIMIT prompts consists of 2 dialects read by 630 speakers, 450 phonetically compact sentences and 1890 phonetically diverse sentences.

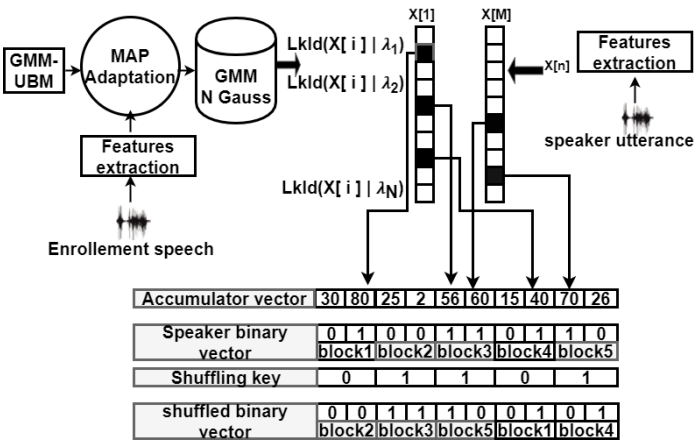

Fig. 2. Cancelable speaker verification template steps

The **RSR2015** database comprises recording from 143 female and 157 male speakers in 9 distinct sessions using mobile devices. The database is divided into three parts according to the lexical content. Part 1 is dedicated to text-dependent scenario where each speaker records 30 sentences per session selected from the TIMIT database. Part2 consists of command and control sentences while Part3 is dedicated for text-prompted speaker verification. In this paper, only the Part 1 of the database is used. In order to have comparable results for our experiments, we apply the proposed evaluation protocol from [21].

Part 1 is divided in disjoint gender-dependent subsets: background, development and evaluations. To train speaker models, for each sentence three repetitions from different sessions were used. The other sessions were used for testing. Part 1 of RSR2015 defines four trials to evaluate text-dependent speaker verification system: **target-correct (tar-c)** where the target speaker pronounces the expected pass-phrase, **target-wrong (tar-w)** where the target speaker pronounces a wrong pass-phrase (a phrase that is different from the enrollment one), **impostor-correct (imp-c)** where an impostor speaker pronounces the expected pass-phrase and **impostor-wrong (imp-w)** where an impostor speaker pronounces a wrong pass-phrase (a phrase that is different from the enrollment one).

Target correct trials are considered as a genuine trials, while the others are considered as impostor trials. The impostor correct tests are more challenging, as the impostor pronounces the expected pass-phrase that is used to train the target speaker model. Based on this protocol, the number of trials for each case is reported in Table I. For impostor wrong test (imp-w), we limit the number of trials to 500.000.

B. Experimental Settings

The feature extraction component is common for the proposed cancelable system and the GMM/UBM system using the MSR Identity Toolbox [22]. The feature vector is composed of 20 MFCC coefficients with their first and second derivative coefficients and the log energy, leading to a vector with a dimension of 63 features.

TABLE I
NUMBER OF TRIALS ON THE PART I OF THE RSR2015 DATABASE FOR BOTH MALE AND FEMALE PROTOCOLS USED IN OUR EVALUATION

Trial	Female		Male	
	developement	*evaluation*	*developement*	*evaluation*
tar-c	8.419	8.631	8.931	10.244
tar-w	244.123	250.299	259.001	297.076
imp-c	387.230	414.249	437.631	573.664
imp-w	500.000	500.000	500.000	500.000

As described in section III, to extract speaker binary vector we need to train a GMM model for each speaker. For this a UBM gender dependent models are trained with 1024 Gaussians. Each GMM model is trained by adapting the mean of the UBM to the enrollment sentence, using MAP criterion. As mentioned in the previous section, in Part 1 of RSR2015, 3 utterances from different sessions are used to create GMM model. In the binarization step, the speaker model GMM will be presented in binary format to transform it in the cancelable template. During enrollment, the same three sentences $enrollspeech_g$ used to create a specific speaker GMM model GMM_g are employed to extract three binary vectors throughout this GMM model as described in section III. Multiplication of these vectors returns a speaker binary reference vector with 1024 bits length. This represents the speaker template which is transformed with the speaker's shuffling key $shuffling - key_g$. Then, the shuffled binary template is stored in the database and the shuffling key is deleted. During verification, the user presents its biometric features with its shuffling key. To compute genuine score of the client g, the system generates the shuffled binary vector and compares it with the enrollment shuffled binary reference using Hamming distance. For impostor trial I, an adversary will try to extract the shuffled binary vector from its test utterance $testspeech_I$, model GMM of genuine GMM_g and its own shuffling key $shuffling - key_I$. The impostor score is a Hamming distance between two binary vectors from different speakers extracted from the same model GMM and combining with different shuffling keys.

C. Binary Gaussian Model Analysis

In order to test the effect of the binarization, we start with setting the parameters $\theta 1$ and $\theta 2$. The model GMM is trained with 512 Gaussians using TIMIT database. Speaker binary vector is extracted leading to 512 bits. In this experience we search the optimum value for $\theta 1$ and $\theta 2$ that minimizes the equal error rate on the development data. Figure 3 shows the EER distribution between genuine-genuine and genuine-impostor scores with respect to $\theta 2$. For $\theta 2 < 20\%$ binary vector cannot discriminate between speakers, because the most selected positions coincide to Gaussians that model noisy acoustic frames. Also, for $\theta 2 > 40\%$ the discrimination of audio binary codes is complicated. We note the optimum value for $\theta 2$ is 30%. Running a similar experiment for $\theta 1$, we observe that $\theta 2$ = 2% minimizes the EER.

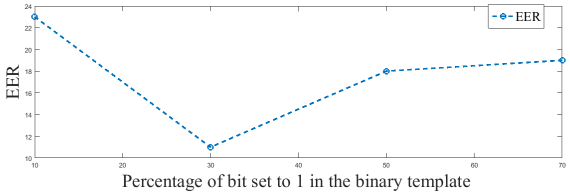

Fig. 3. EER distribution for speaker binary template according to parameter $\theta 2$ on the develoment TIMIT database

D. Performance Analysis on RSR2015

For objective comparison, the verification performance of the proposed cancelable system is compared with the baseline UBM-GMM. To measure the performance, the Equal Error Rate (EER) is computed for both the development and evaluation parts. Table II and III report respectively the EER on the development and evaluation subset. As seen the system based on the shuffling scheme outperforms the UBM-GMM. In the evaluation subset (female and male), a clear improvement in the challenging trial impostor correct is shown. For female tests, the EER of the baseline system is 1.8% which reduces to 0.01% when shuffling scheme is applied. Also for the male tests, the EER of the baseline is 0.43% which reduces to 0.01% with the proposed cancelable binary system. Figure 4 shows the distribution of the target correct and impostor correct Hamming distance for female test on the evaluation subset before and after shuffling. The shuffling process increases the impostor Hamming distances while the genuine scores remain unchanged. The mean of the impostor Hamming distance distribution of the baseline system shifts from 0.27 to 0.42 when the shuffling scheme is applied. This reduces the overlap between genuine and impostor distributions which improves the discrimination capacity of the system and thereby leads to a better verification accuracy.

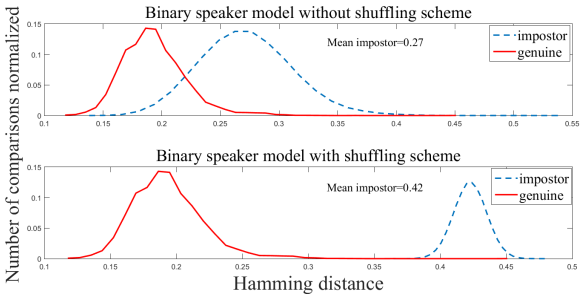

Fig. 4. Normalized Hamming distance distributions for target correct and impostor-correct comparisons on the female evaluation subset of Part1 RSR2015 database

E. Revocability Analysis on RSR2015

The revocability is evaluated by comparing protected templates generated from distinct shuffling keys. Speaker specific binary vector is transformed using 300000 different shuffling keys to generate 300000 protected templates. The first shuffled binary vector is compared with the others to

TABLE II

PERFORMANCE OF THE PROPOSED CANCELABLE SPEAKER VERIFICATION SYSTEM AND THE UBM-GMM BASELINE SYSTEM ON THE DEVELOPMENT SUBSET OF RSR2015 IN TERMS OF EER FOR THE TARGET-CORRECT, IMPOSTOR CORRECT AND IMPOSTOR WRONG

EER%	**female**			**male**		
	UBM-GMM	***binary template***	***cancelable system***	***UBM-GMM***	***binary template***	***cancelable system***
tar-correct/imp-correct	3.31	20	3	3.5	16.05	1.32
tar-correct/imp-wrong	0.25	8.05	2.9	0.9	7.31	2.18

TABLE III

PERFORMANCE OF THE PROPOSED CANCELABLE SPEAKER VERIFICATION SYSTEM AND THE UBM-GMM BASELINE SYSTEM ON THE EVALUATION SUBSET OF RSR2015 IN TERMS OF EER FOR THE TARGET-CORRECT, IMPOSTOR CORRECT AND IMPOSTOR WRONG

EER%	**female**			**male**		
	UBM-GMM	***binary template***	***cancelable system***	***UBM-GMM***	***binary template***	***cancelable system***
tar-correct/imp-correct	1.98	10.25	0.01	2.2	22	0.6
tar-correct/imp-wrong	0.43	2.62	0.01	3.1	14.6	0.51

compute the pseudo-impostor scores. The process is repeated with 30 different users. From Figure 5, we can see that the pseudo-impostor score distribution resembles the impostor distribution. This result validates that the generated shuffling binary vectors are indistinguishable to each other, although they are generated from the same binary code (same user). As a result, in case of compromise, a cancellation is possible and a new template can be generated via another shuffling key.

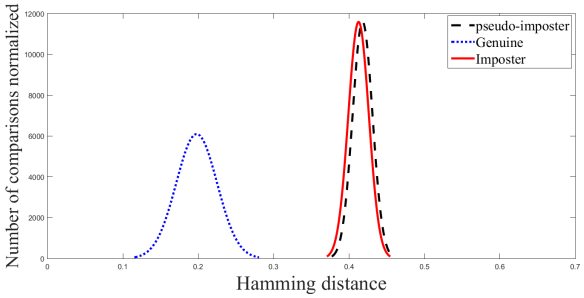

Fig. 5. Distribution of genuine, impostor and pseudo-impostor scores on the female evaluation subset of Part1 RSR2015 database

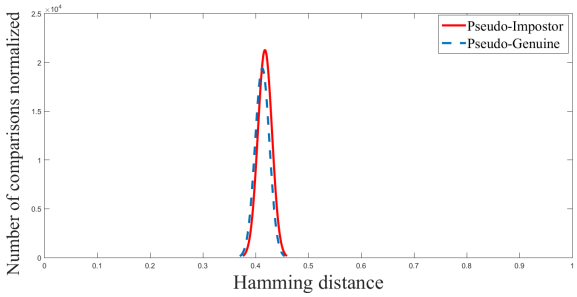

Fig. 6. Distribution of pseudo-genuine scores and pseudo-impostor scores on the female evaluation subset of Part1 RSR2015 database

TABLE IV

SECURITY ANALYSIS OF THE PROPOSED CANCELABLE SYSTEM ON THE FEMALE EVALUATION SUBSET OF PART1 RSR2015 IN TERMS OF EER %

System	EER%
baseline UBM-GMM	0.43
binary template	2.62
cancelable template	0.01
stolen biometric	0
stolen key	2.62

F. Unlinkability Analysis

The unlinkability is evaluated by introducing the pseudo genuine scores. Pseudo genuine scores are computed by matching different binary codes generated from the same user with different shuffling keys. The overlapping of pseudo impostor scores and pseudo-genuine scores will indicate whether the shuffling binary codes generated from the same user or from another are different. As shown in Figure 6 there is a large overlap between the pseudo-impostor distribution and pseudo-genuine distribution. Hence, this suggests that the shuffling transformation is able to fulfill the unlinkability property.

G. Security Analysis

1) Brute force attack: A brute force attack consists of an adversary trying to guess the correct template to access as genuine user. In the proposed system the similarity score is computed based on the similarity between the enrolled and tested shuffled binary vector. If the similarity score exceeds the threshold, the user will be deemed as the legitimate user. In the proposed system the threshold is 0.012. If the adversary wants to guess the correct value of shuffled binary vector with length of 1024 bits, the guessing complexity is $2^{1024*(1-0.012)}$ attempts.

2) Stolen Biometric Feature: In the stolen biometric scenario, we suppose that the biometric template for all the speakers is compromised. In such case, an adversary provides the stolen biometric template with an incorrect shuffling key to gain access to the system. As seen in Figure 7 the distribution for stolen biometric scenario is overlapped with the impostor distribution. This demonstrates that two binary vectors of the same speaker shuffled using different shuffling keys appear to be generated from two distinct speakers. As

shown in Table IV, the EER in stolen biometric case is 0%. Thus, if an impostor has the stolen biometric template of a legitimate person to get verified, the system still resists.

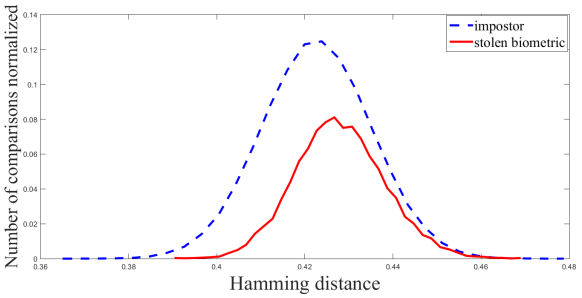

Fig. 7. Distribution of impostor scores and stolen biometric scores on the female evaluation subset of Part1 RSR2015

3) Stolen key: In the stolen key scenario, we study another privacy threat when the shuffling key of all the speakers is stolen. As shown in Table IV, the performance is equal to the EER of the Gaussian binary vector. Indeed, the dissimilarity between two shuffled binary vectors increases only if they are shuffled with different keys. Using the same key to transform different vectors, the dissimilarity scores do not change. Thus, in this case, the performance is exactly the same as that for the baseline Gaussian binary vectors.

V. Conclusion

A cancellable biometric system based on binary Gaussian model with shuffling key has been proposed for speaker verification. Experimental results show that the proposed cancelable system improves the verification performance and satisfies the evaluation criteria of the template protection, as revocability and unlinkability. Due to the shuffling scheme, the shuffled speaker template has the desired property of cancelability. The proposed system can generate different templates for various applications using the same biometric features which conserves privacy. If the stored shuffled template is compromised, the administrator can cancel the old template and issue a new one by changing the shuffling key. In addition, comparing with existing work on voice biometric protection, the proposed system achieves a significant improvement in biometric performance that assures EER of 0.01% and even if the biometric feature is stolen, the ERR of the proposed system still better than the baseline biometric system.

Acknowledgment

This work is partially supported by the SpeechXRays project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 653586.

References

[1] F. Bimbot, J.-F. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau, S. Meignier, T. Merlin, J. Ortega-García, D. Petrovska-Delacrétaz, and D. A. Reynolds, "A tutorial on text-independent speaker verification," *EURASIP journal on applied signal processing*, vol. 2004, pp. 430–451, 2004.

[2] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," *IEEE security & privacy*, vol. 99, no. 2, pp. 33–42, 2003.

[3] S. G. Kanade, D. Petrovska-Delacrétaz, and B. Dorizzi, "Cancelable biometrics for better security and privacy in biometric systems," in *International Conference on Advances in Computing and Communications*. Springer, 2011, pp. 20–34.

[4] A. K. Jain, A. Ross, and U. Uludag, "Biometric template security: Challenges and solutions," in *Signal Processing Conference, 2005 13th European*. IEEE, 2005, pp. 1–4.

[5] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," *EURASIP Journal on Information Security*, vol. 2011, no. 1, p. 3, 2011.

[6] K. Nandakumar and A. K. Jain, "Biometric template protection: Bridging the performance gap between theory and practice," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 88–100, 2015.

[7] "ISO/IEC 24745:2011. Information technology — Security techniques — Biometric information protection," International Organization for Standardization, Standard, 2011.

[8] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

[9] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, "Generating cancelable fingerprint templates," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 561–572, 2007.

[10] A. B. J. Teoh and L.-Y. Chong, "Secure speech template protection in speaker verification system," *Speech communication*, vol. 52, no. 2, pp. 150–163, 2010.

[11] A. Juels and M. Sudan, "A fuzzy vault scheme," *Designs, Codes and Cryptography*, vol. 38, no. 2, pp. 237–257, 2006.

[12] W. Xu and M. Cheng, "Cancelable voiceprint template based on chaff-points-mixture method," in *Computational Intelligence and Security, 2008. CIS'08. International Conference On*, vol. 2. IEEE, 2008, pp. 263–266.

[13] M. Paulini, C. Rathgeb, A. Nautsch, H. Reichau, H. Reininger, and C. Busch, "Multi-bit allocation: Preparing voice biometrics for template protection," *Odyssey 2016*, pp. 291–296, 2016.

[14] S. Billeb, C. Rathgeb, H. Reininger, K. Kasper, and C. Busch, "Biometric template protection for speaker recognition based on universal background models," *IET Biometrics*, vol. 4, no. 2, pp. 116–126, 2015.

[15] J.-F. Bonastre, P.-M. Bousquet, D. Matrouf, and X. Anguera, "Discriminant binary data representation for speaker recognition," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 5284–5287.

[16] J.-F. Bonastre, X. Anguera, G. H. Sierra, and P.-M. Bousquet, "Speaker modeling using local binary decisions," in *Twelfth Annual Conference of the International Speech Communication Association*, 2011.

[17] X. Anguera and J.-F. Bonastre, "A novel speaker binary key derived from anchor models," in *Eleventh Annual Conference of the International Speech Communication Association*, 2010.

[18] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted gaussian mixture models," *Digital signal processing*, vol. 10, no. 1-3, pp. 19–41, 2000.

[19] S. Kanade, D. Camara, E. Krichen, D. Petrovska-Delacrétaz, and B. Dorizzi, "Three factor scheme for biometric-based cryptographic key regeneration using iris," in *Biometrics Symposium, 2008. BSYM'08*. IEEE, 2008, pp. 59–64.

[20] P. A. Keating, D. Byrd, E. Flemming, and Y. Todaka, "Phonetic analyses of word and segment variation using the timit corpus of american english," *Speech Communication*, vol. 14, no. 2, pp. 131–142, 1994.

[21] A. Larcher, K. A. Lee, B. Ma, and H. Li, "Text-dependent speaker verification: Classifiers, databases and rsr2015," *Speech Communication*, vol. 60, pp. 56–77, 2014.

[22] S. O. Sadjadi, M. Slaney, and L. Heck, "Msr identity toolbox v1. 0: A matlab toolbox for speaker-recognition research," *Speech and Language Processing Technical Committee Newsletter*, vol. 1, no. 4, pp. 1–32, 2013.