Realization of Correlation Attack Against Fuzzy Vault Scheme

Alisher Kholmatov and Berrin Yanikoglu

Faculty of Engineering & Natural Sciences
Sabanci University, Istanbul, 34956, Turkey

ABSTRACT

User privacy and template protection are major concerns impeding deployment of the biometrics to the general public, since once compromised, most of the biometric traits can not be canceled or reissued.

Fuzzy vault scheme,[1] emerged from recent research efforts to alleviate the problem. It makes use of scrambling redundant and a user dependent data such that during verification genuine user can easily identify substantial amount of his/her data (thus authenticate him/her self), which is computationally infeasible for a forger. However it was recently claimed that the fuzzy vault is susceptible against correlation based attacks,[2] where it is assumed that an attacker could intercepted a number of fuzzy vaults issued for the same biometric trait.

In this work we implement correlation based attacks against the fuzzy vault scheme and quantify our results using database of 400 fingerprints. As a matter of fact, we could correlate 59% of vaults approving the claim of fuzzy vault's vulnerability against the correlation attack.

Keywords: Fuzzy, Vault, Attack, Template, Protection, Biometrics, Privacy

1. INTRODUCTION

As biometrics are gaining popularity, there is increased concern over the loss of privacy and potential misuse of biometric data held in central repositories. Biometric data which can uniquely identify a person (e.g. fingerprints, iris patterns) can be used to track individuals, linking many separate databases (where the person has been, what he has purchased etc.). There is also fear that the central databases can be used for unintended purposes.[3] For instance, latent fingerprints can be used to search for information about a person in a central database, if such databases are compromised. The association of fingerprints with criminals raise further concerns for fingerprint databases in particular. Similarly, biometric data may reveal certain rare health problems,[4] which raises concern about possible discriminatory uses of central databases. On the other hand, the alternative suggestion of keeping biometric data in smart cards does not solve the problem, since forgers can always claim that their card is broken to avoid biometric verification altogether.

Ratha et al. suggest[5] and implements[6] a framework of cancelable biometrics, where a biometric data undergoes a predefined non-invertible distortion during both enrollment and verification phases; if the transformed biometric is compromised, the user is re-enrolled to the system using a new transformation. Likewise, different applications are also expected to use different transformations for the same user. Although this framework hides original (undistorted) biometric and enables revocation of a (transformed) biometric, it introduces the management of transform databases, and still requires registration of reference points.

Yanikoglu et al.[7] propose a biometric authentication framework which uses two separate biometric features combined to obtain a non-unique identifier of the individual, in order to address privacy concerns. As a particular example, they demonstrate a fingerprint verification system that uses two separate fingerprints of the same individual. A combined biometric ID composed of two fingerprints is stored in the central database and imprints from both fingers are required in the verification process, lowering the risk of misuse and privacy loss. They show that the system is successful in verifying a persons identity given both fingerprints, while searching the

Further author information: (Send correspondence to Alisher Kholmatov)
Alisher Kholmatov: e-mail: alisher@su.sabanciuniv.edu, web: http://students.sabanciuniv.edu/~alisher
Berrin Yanikoglu: e-mail: berrin@sabanciuniv.edu, web: http://people.sabanciuniv.edu/berrin

combined fingerprint database using a single fingerprint, is impractical. Although successful with fingerprints it is not straightforward how that framework could be extended to other biometric modalities.

Jules et al. propose the *fuzzy vault* scheme[1] and described how it can be used to encapsulate and release an encryption key using one's biometrics. A secret (cryptographic key) is *locked* using a biometric data of a person and a large amount of redundant data that conceal the biometric, such that only someone who possesses a substantial amount of the locking elements (e.g. another reading of the same biometric) would be able to decrypt the secret. Fuzzy vault construct is an example of recent research which focus on combining cryptography and biometrics to take advantage of the benefits of both fields[1, 8–11]: while biometrics provide non-repudiation and convenience, traditional cryptography provides adjustable levels of security and can be used not just for authentication, but also for encryption.

On the other hand, not much research work is available on formal security analysis of such privacy preserving schemes. In one of the few, Scheirer et al.[2] sugest a number of prominent attacks and elaborates on their impact on fuzzy vault & biometric encryption schemes. In this study we aimed to empirically quantify vulnerability of fuzzy vault scheme against correlation based attack.

2. RELATED WORK

Fuzzy Vault

Jules and Sudan proposed a scheme called *fuzzy vault*, which they call an error tolerant encryption operation.[1] Fuzzy vault scheme provides a framework to encrypt ("lock") some secret value (eg. cryptographic key) using an unordered set of locking elements as a key, such that some one who possesses a substantial amount of the locking elements will be able to decrypt the secret. It is based on the difficulty of the polynomial reconstruction problem. The encoding and decoding are done as follows:

Assume that Alice wants to secure her cryptographic key S (a random bit stream) using an arbitrary set of elements A. She selects a polynomial $P(x)$ of degree D and encodes S into the polynomial's coefficients. Encoding can be achieved by slicing S into non-overlapping bit chunks and then mapping these onto the coefficients. The mapping must be invertible meaning that the coefficients can be unambiguously mapped back to the corresponding bit chunks, which when concatenated will reconstruct the S. Then, Alice evaluates the polynomial at each element of her set A and stores these evaluation pairs into the set G, where $G = \{(a_1, P(a_1)), (a_2, P(a_2)), ..., (a_N, P(a_N))\}$, $a_i \in A$ and $|A| = N$. Finally, she generates a random set R of pairs such that none of the pairs in that set lie on the polynomial; and she merges the sets G and R into a final set, to obtain the vault, which she then makes public. Note that within the vault, it is not known which points belong to G and which ones belong to R. All the steps required to lock a secret in the Fuzzy Vault are graphically represented in Figure 1.

Now suppose that Bob has his own set of elements B and he wants to find out ("unlock") Alice's secret locked in the vault. He will be able to do so only if his set B largely overlaps with Alice's A, so as to identify a substantial number of the pairs that lie on the polynomial, from the vault. Given at least $D + 1$ pairs that lie on the polynomial, he applies one of the known polynomial reconstruction techniques (eg. Lagrange interpolating polynomial) to reconstruct the polynomial and thus extracts her secret S. Notice that if Bob does not know which of the points of the vault lie on the polynomial, it should be computationally infeasible for him to unlock the vault.

Whereas perturbation of a single bit in a key of a classical cryptosystem (eg. AES, RSA) hinders decryption completely, the fuzzy vault allows for some minor differences between the encryption & decryption keys, here the unordered sets used to lock & unlock the vault. This fuzziness is necessary for use with biometrics, since different measurements of the same biometric often result in quite different signals, due to noise in the measurement or non-linear distortions. Furthermore, for most of the known biometric signals, it is hard to establish a consistent ordering within the measured features. For instance two impressions of the same fingerprint can have substantial distortion and the number of features may vary between the two impressions. require ordering fuzziness, it is not straightforward how to implement the fuzzy vault using biometric data, due to the difficulty of matching the query and template biometric signals (i.e. locking and unlocking sets, respectively) especially within the presence of random data (the chaff points).

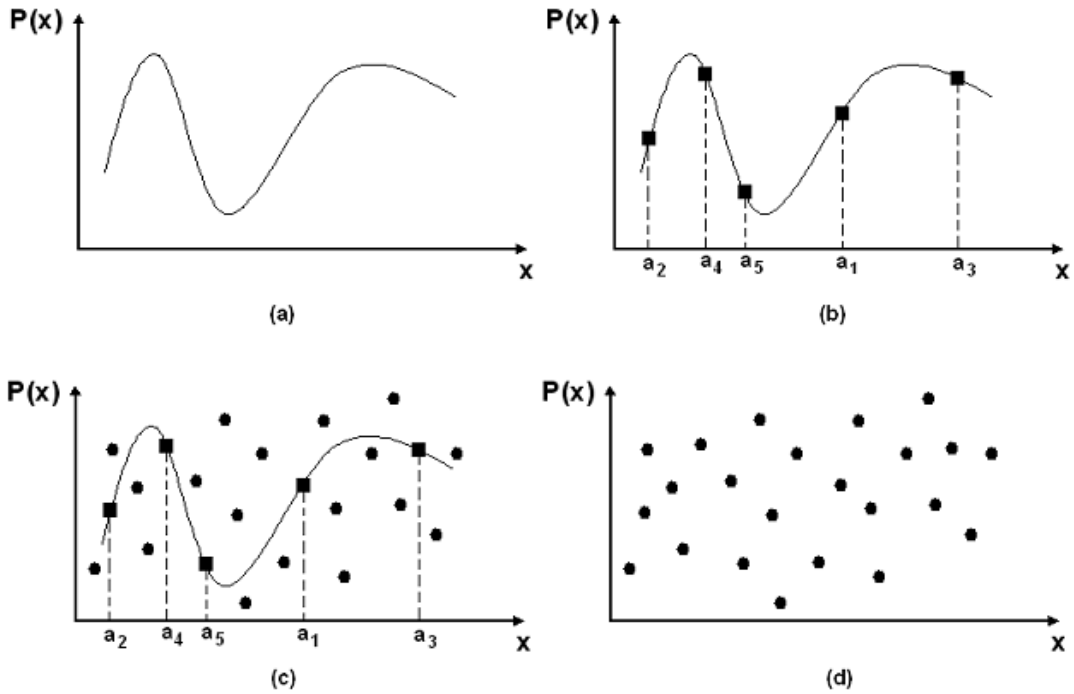

Figure 1. Vault *Locking* phase: (a) Create a polynomial by encoding the *Secret* as its coefficients. (b) Project genuine features onto the polynomial: a_i represents the subject's i'th feature. (c) Randomly create chaff points (represented by small black circles) and add to the Vault. (d) Final appearance of the Vault, as stored to the system database.

Fuzzy Vault with Fingerprints

Uludag et al.[12] demonstrated a preliminary implementation of the fuzzy vault scheme using fingerprints. Yang and Verbauwhede[13] also implemented the fuzzy vault with fingerprints, but they made the assumption that rotation & translation invariant features can be reliably extracted from minutiae, which is difficult in practice. Furthermore, they store reference minutia point along with the vault, which may also leak some information. We will review the system by Uludag et al. as it relates the most to our proposed scheme.

Minutia points of template & query fingerprints were used as locking & unlocking sets, respectively, to lock a 128-bit long data (S) which forms the cryptographic key. More precisely, the values obtained by concatenation of the corresponding x & y coordinates of minutiae points were used as set elements. To make sure that the desired S was unlocked from the vault through an error-prone process, cyclic redundancy check bits (16 bits) were concatenated to S. Then, S, together with its check bits, was divided into non-overlapping chunks (16 bits each), giving the coefficients, of an 8th degree polynomial. To lock the secret, template minutiae set was projected onto this polynomial and random chaff points not lying on the polynomial are added, to form the vault. Based on their empirical estimations, they used only 18 minutia points and 200 chaff points.

To unlock the secret, i.e. reconstruct S, they first match the query minutia set with the abscissa part of the vault and identify candidate points lying on the polynomial. Since $D+1$ points are required to reconstruct a polynomial of degree D, all possible 9 point combinations of the candidate set are tried, to find the one with the correct check bits. S is successfully unlocked when the check bits verify. Authors report a 79% of correct reconstruction rate with 0% false accept rate.

To bypass the problem of matching the minutiae points and finding an upperbound for the performance of the scheme, the authors have used a fingerprint database where minutia points and the correspondence between template & query fingerprints were established by an expert. During their experiments, the minutiae sets of

mating fingerprints were pre-aligned (i.e. rotated & translated) according to the established correspondence, and used as such. However, later[14] authors extracted distinguishing points from fingerprint's ridge curves and used these as a helper data for vault & test fingerprint alignment.

Attacks on Fuzzy Vault

Scheirer et al.,[2] inspired by classical cryptographic security analyses, suggests a number of attack scenarios on biometric privacy preserving schemes with a special focus on fuzzy vault[1] & biometric encryption[15] schemes. They classify their attacks into 3 groups: namely i) attacks via record multiplicity, ii) stolen key-inversion attack, and iii) blended substitution attack.

In the case of attacks via record multiplicity, it is assumed that an attacker could somehow intercept multiple enrollments/encodings of a particular privacy preserving scheme, created using the same biometric data (eg. a number of fuzzy vaults created using different imprints of the same fingerprint, using different chaff points). It is suggested that, it may be possible to correlate these records and retrieve the biometric trait itself or link databases.

Stolen key-inversion attack presumes that an attacker could get access (eg. by means of social engineering, weak coupling between modules of a security system, etc.) to a secret key released upon positive authentication to a system. Then, utilizing the key the attacker could reconstruct the biometric trait. In the case of fuzzy vault, with the key in hand, it is a straightforward task to identify genuine & chaff points by first constructing the polynomial and then verifying weather or not a corresponding point lies on the polynomial.

Finally, the blended substitution attack considers the situation where a malicious attacker injects his own data into someone's template such that both genuine & malicious users will be positively authenticated against the same enrollment record. In the case of the fuzzy vault scheme, the attacker would insert his own minutiae points, possibly using his own secret (polynomial). When he presents his biometric, the original user's minutiae points will act as chaff points and the forger will get authenticated. While the implementation of this attack may not be straightforward, it is apparent that the fuzzy vault may be susceptible for this type of attack. In fact, Kholmatov et al.[16] exploited this additive property of the fuzzy vault scheme to implement a biometric based secret sharing.

In their paper, Scheirer et al. only made theoretical claims about the weaknesses of these systems against these 3 type of attacks; however, they have not provided any implementation to support their claims. In this paper, we tried to address this issue and analyze the weaknesses of the fuzzy vault scheme against the proposed attacks.

3. IMPLEMENTATION OF CORRELATION BASED ATTACKS

In this study we aimed to realize and empirically quantify attacks against the fuzzy vault scheme. We had previously implemented the scheme using fingerprint minutiae[16] as proposed by,[12] although other implementations could also be considered. Due to the fact that the impact of the stolen key-inversion is obvious, and the one for the blended substitution attack is relatively straightforward, we focus only on the attack via record multiplicity.

Three different questions need to be answered: first, given the fact that an attacker obtains two fuzzy vaults which are locked using the same biometric trait but different chaff points and secrets (i.e. polynomial), how computationally easy is it to reconstruct the secrets encapsulated in the corresponding vaults? Second, given a fuzzy vault created using a biometric trait of an unknown person and a database of vaults linked to their corresponding identities, what is the chance of correctly identifying the unknown person? Finally, given two databases of vaults, how likely is it to correctly link corresponding vaults? Note that the second scenario is a sub-task of the third case.

We have collected a database of 200 different fingers using an optical sensor, where for each finger 2 different impressions were obtained (400 fingerprints in total). We have manually labeled the minutiae points of each fingerprint and created corresponding fuzzy vaults using different chaff points and secrets for each impression. The fuzzy vaults were locked using 200 chaff points and polynomials of degree 8, while all available minutiae points were used without a restriction.

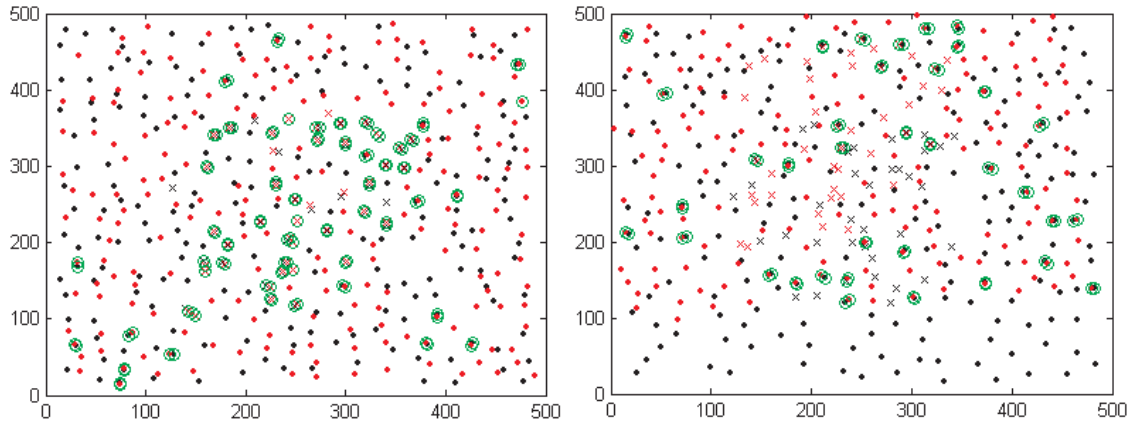

Figure 2. This figure depicts alignments of two vaults created using the same fingerprint (on the left) and different fingerprints (on the right). Dot points (red & black) and crosses (red & black) identify chaff and minutiae points of corresponding vaults, respectively, where matching points are encircled.

Correlating two fuzzy vaults

In order to answer the first question, we correlated the fuzzy vaults of corresponding fingerprint impressions using exhaustive matching. The exhaustive matching seeks for the best alignment (in the sense of number of matching points) between the vaults, while exploring all possible affine transformations (rotations & translations only) between the vaults. After the alignment we calculated the proportion of matching genuine & chaff points. The vaults are considered to be successfully correlated if i) number of identified genuine points (G) is sufficient to decode the polynomial of degree D (i.e. $G \geq D + 1$) and ii) the total number of correlated points (N) is not too large, so that brute force attack (i.e. trying all possible combinations of D+1 out of N points) to reconstruct the polynomial is computationally feasible. The average number (A) of such combinations the attacker would try is formulated as follows:

$$A = \frac{\binom{N}{D+1}}{\binom{G}{D+1}}$$

On average for matching vaults, the exhaustive matcher identified 43 point correspondence (i.e. N), where 22 out of these points were genuine (i.e. G). Based on above formulation, 1133 attempts on average are required to reconstruct a vault, which takes a couple of seconds for a modern desktop personal computer (PC). In each attempt, one can be certain of having decoded the correct secret stored in the vault thanks to the error detection scheme of the vault. The whole process takes approximately 50 seconds for a Matlab implementation (exhaustive matching and decoding attempts) on a 64-bit Xeon 3GHz powered PC.

Using this method, we could successfully reconstruct 59% of the corresponding fuzzy vault pairs (i.e. 118 out 200) in our fingerprint database. In other words, given two fuzzy vaults constructed with the same fingerprint (but different impressions), one can be extract the fingerprint data embedded in the fuzzy vault, without any ambiguity, with a high probability. Figure 2 depicts alignments of two vaults created using the same (on the left) and different (on the right) fingerprints. As can be seen, genuine points greatly align in vaults created using the same fingerprint, which is not in the other case.

We could also try to use alignments where number of correlated points exceed certain threshold, instead of using only the best alignment in order to increase the number of successful attacks (successfully decoded secrets), though we didn't try this.

Correlating two databases

The second question is answered by correlating each fuzzy vault with the rest of the vaults (1 to 399 comparisons for our database). One may expect that the matching vault (created using the second impression of the fingerprint) would have the highest number of matching points among all the vaults compared. For that reason, we aligned the query vault with all the remaining vaults in the database (one of them being the matching vault) and picked the one with the highest number of matching points (i.e. N), using the exhaustive match described above. Then, we tried to decode the secret embedded in the vault by trying all possible combinations of D+1 matching points, where D is the degree of the polynomial. Due to the time consuming nature of this attack, we tried to match only 10 fuzzy vaults against all the other vaults and observed that 4 out of 10 matching vaults were identified.

When we tried to decode not only the best aligned vault, but the top-10 best aligned vaults, the success rate increased to 6 out of 10. In other words, out of the 10 vaults matched against the whole database, 6 were successfully matched to their corresponding vaults.

This experiment prove the claim that the fuzzy vault scheme without further security measures is not sufficient to sustain users' privacy. (Note to reviewers: full attack results using all 400 vaults will be presented in the final version of the paper.)

4. SUMMARY AND CONCLUSION

The fuzzy vault scheme emerged as a promising solution for preserving privacy in biometric based authentication systems. In this paper, we studied the effectiveness of a correlation based attack against the fuzzy vault scheme. We empirically quantified the complexity required for an attacker to compromise the scheme, with a high degree of success: we were able to retrieve the secrets from 59% of corresponding vault pairs supporting the claim that the fuzzy vault scheme is indeed vulnerable against such attacks.

ACKNOWLEDGMENTS

This work is supported by TÜBİTAK (The Scientific and Technical Research Council of Turkey), under project number 105E165.

REFERENCES

1. A. Juels and M. Sudan, "A fuzzy vault scheme," *IEEE International Symposium on Information Theory* , p. 408, 2002.
2. W. J. Scheirer and T. E. Boult, "Cracking fuzzy vaults and biometric encryption," *Univ. of Colorado at Colorado Springs, Tech. Rep.* , February 2007.
3. G. Tomko., "Biometrics as a privacy-enhancing technology: Friend or foe of privacy?," *In Privacy Laws & Business 9th Privacy Commissioners/Data Protection Authorities Workshop* , 1998.
4. W. H. I. McLean, "Genetic disorders of palm skin and nail," *Journal of Anatomy* **202(1)**, p. 133133, 2003.
5. N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Syst. J.* **40**(3), pp. 614–634, 2001.
6. "Generating cancelable fingerprint templates," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), pp. 561–572, 2007. Nalini K. Ratha and Sharat Chikkerur and Jonathan H. Connell and Ruud M. Bolle.
7. B. Yanikoglu and A. Kholmatov, "Combining multiple biometrics to protect privacy," in *ICPR-BCTP Workshop, Cambridge, England*, August 2004.
8. P. Tuyls, E. Verbitskiy, T. Ignatenko, D. Denteneer, and T. Akkermans, "Privacy protected biometric templates: Acoustic ear identification," *Proceedings of SPIE: Biometric Technology for Human Identification* **Vol. 5404**, pp. 176–182, 2004.
9. G. Davida, Y. Frankel, and B. Matt, "On enabling secure applications through on-line biometric identification," *In IEEE Symposium on Privacy and Security* , p. 408, 1998.

10. C. Soutar, D. Roberge, S. Stojanov, R. Gilroy, and B. V. Kumar, "Biometric encryption using image processing," *In Proc. SPIE, Optical Security and Counterfeit Deterrence Techniques II* **Vol. 3314**, pp. 178–188, 1998.
11. J. Linnartz and P. Tuyls, "New shielding functions to enhance privacy and prevent misuse of biometric templates," *Proceeding of AVBPA* (LNCS 2688), pp. 393–402, 2003.
12. U. Uludag, S. Pankanti, and A. Jain., "Fuzzy vault for fingerprints," *Proceeding of International Conference on Audio- and Video-Based Biometric Person Authentication* , pp. 310–319, 2005.
13. S. Yang and I. Verbauwhede, "Secure fuzzy vault based fingerprint verification system," *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on* , pp. 577–581, 2004.
14. U. Uludag and A. K. Jain, "Securing fingerprint template: fuzzy vault with helper data," in *Proc. IEEE Workshop on Privacy Research In Vision*, June 22, 2006.
15. C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and V. Kumar, *Biometric Encryption, Chapter 22 of the ICSA Guide to Cryptography*, R.K. Nichols, McGraw-Hill, Ed. 1999.
16. A. Kholmatov, B. A. Yanikoglu, E. Savas, and A. Levi, "Secret sharing using biometric traits," in *Biometric Technology For Human Identification III, Proceedings of SPIE*, **6202**, 18 April, 2006.