Student Name: Panchal Gaurangkumar Rameshbhai

Roll No: 12IT91Q01

Title: Fingerprint Biometric-Based Approaches for Remote User Authentication Schemes

Response to the comments of Foreign Examiner (Prof. Alex Kot, NTU Singapore)

Weakness of the Thesis (of 12IT91Q01))

1. Overhead, vulnerabilities and challenges were mentioned for crypto-biometric system, user authentication at remote location and biometric-based multiparty authentication system, but references should be cited accordingly so that readers know the weaknesses and the limitations of the existing techniques in Chapter 1.

   Response:
   We have mentioned all these in Chapter 3-5, where we discuss three research problems namely crypto-biometric system, remote authentication and multiparty authentication system. We feel that it is more appropriate to cover overheads/vulnerabilities/challenges in the respective chapters instead of in Chapter 1 as suggested by the examiner.

   [Please see Chapter 1, Page 13-15 in the revised thesis.]

2. The research objectives are listed, but what are the current works that the author tries to improve. References are needed in Chapter 1.

   Response:
   In Chapter 2, we have specifically mentioned current state of the art vis-a-vis our research objectives. There, we clearly mentioned "Summary of Observations" and "Table of Summary" stating limitations/issues in the existing work.

3. In the overview of the research work, the author should be more specific to state what are the novelties and contributions of his work.
   I realize that the limitations and vulnerabilities of the existing work are described in the end of Chapter 1. However, it is desirable to use a few sentences to highlight them in the beginning also.

   Response:
   As per the suggestion of the examiner, we have mentioned the "research contribution" in Chapter 1.

   [Please see Chapter 1, Page 13-15 in the revised thesis.]

4. Some statements are over claimed. For example, the distinctiveness of using fingerprint to generate the bio-crypto has not be resolved yet if the fingerprint is not perfectly captured.

   Response:
   Such over claimed statements have been removed.

Questions to be asked during Thesis Defense

1. It seems that an assumption is made such that the reference fingerprint and the testing fingerprint will have the same number of core/delta points detected or not being detected. It is possible that multiple core/delta points are detected in one device during the encryption process and single or no core/delta points are detected from another device during the decryption phase. How to handle the key size under this condition?

   Response:

   If core and/or delta points are not detected.

   It is possible that there are some images, where delta or core point (or both) is (are) missing. In such a situation, the feature sets F_C and/or F_D will not be generated. Then, we use F_A to generate key. We follow the initial key bit generation and expansion-permutation mechanism to ensure the key bits (Section 3.2.1.4).

   If multiple core and/or delta points are detected.

   In case of multiple core points are found, then we consider the all core points based on the increasing order of their x-coordinate value. If the x-coordinate values are same for more than one core points, then we consider that core points using increasing order of their y-coordinate values. We apply our core point based approach for each detected core point of a fingerprint and put the straight lines attributes in the set F_C. Similar mechanism is followed in case of multiple delta points are detected. We use these sets to generate the key bits. We follow the initial key bit generation and expansion-permutation mechanism to ensure the key bits (Section 3.2.1.4).

   In general, during encryption phase, we generate a feature vector (F_1) from the captured fingerprint image. Here, F_1 may contain the biometric information based on core/delta point(s) detected or not being detected. We use F_1 to generate codeword (C_1) and bio-crypto key (K_1). During decryption phase, we generate a feature vector (F_2) from newly captured fingerprint image. Here, F_2 may/may not contain core/delta points' information. Note that F_2 is compared with F_1 (decoded C_1). If user is authenticated then we generate key from F_1. We generate key from codeword, which is obtained from a single instance of fingerprint; hence, different size of keys do not arise.

2. The author mentioned on page 41 that the error correcting capability of Reed-Solomon code is maximum up to 16 errors, according to ref 140. How does this translate to the error tolerance level of the minutiae locations?

   Response:

   We convert each value of feature vector (i.e., length and angle) of minutiae information into the form of codeword. In other words, each value is considered as a message. We create one error in each message to generate a codeword. Thus, the error tolerance level is 1 in our proposed approach. Note that the message length is 1 and error tolerance level 1. During decryption phase, each error is resolved and original message is generated.

   Therefore, error-correcting capability in codeword, no way related to the minutiae locations

3. How to deal with spurious minutiae which will generate wrong keys even with error correction in your data storage security?

Response:
This question does not arise as we do not generate key from two instances of biometric data.

4. Would it be possible to put weightage to minutia that are more prominent and clear so that only the reliable ones are used to generate the key to minimize errors?

   Response:
   It is not necessary to put weightage to minutia that are more prominent and clear. In our work, we generate key from one biometric sample and hence it is not required to consider errors.

5. In section 3.3.3.2, the author uses SNR ration to measure the quality of the fingerprint image. When extracting the minutiae from the fingerprint, binarization step is usually performed. It is well known that SNR is not a good measure for binary images. Can the author explain why SNR is used here? As a matter of fact, test fingerprints that match with the reference fingerprints should be considered as acceptable quality. For example, with such SNR measure, there are 279 good quality fingerprints and 385 bad quality fingerprints in Set B, but most of the fingerprints in the NIST dataset contain acceptable fingerprints in real scenario. As a matter of fact, table 3.3 shows that the similarity in feature vectors is good. Using Good-Bad combination, Table 3.4 shows that there are only 359 successful decryption out of 450 possible combinations. This yields less than 80 successful rate using fingerprints to generate the encryption key. How will the author propose possible improvement?

   Response:
   In our work, SNR is used to measure the quality of fingerprint images. This concept is standard practice and follow in many works.

   To improve the accuracy in Good-Bad combination, we may focus on Region of Interest (ROI) of fingerprint image rather than considering whole image. As the uncertainty of feature existence in the bad fingerprint image is higher which decrease the accuracy.

6. Table 3.10 is used to compare the accuracy against some existing methods. However, these are just the similarity measure of the feature vectors. The ultimate goal in Chapter 3 is to generate a bio-crypto key based on these features to do the encryption. The successfulness of the decryption using the same finger is what it matters. Suggest to remove this table to avoid confusion in performance.

   Response:
   Table 3.10 shows the accuracy measure so far matching is concerned, but not similarity means as the examiner pointed out.

7. For the remote user authentication protocol in Chapter 4, is there any reason why the author does not use core/delta points as one of the reference points for the alignment purpose. This is commonly used in many matching techniques. The suggested approach may not be as reliable as core and delta.

   Response:
   We do not consider core/delta points as one of the reference points for the alignment purpose. This is because, the detection of core/delta points is uncertain in every captured fingerprint image. Hence, the fingerprint image alignment may not be accurate. Hence, we consider the boundary of captured fingerprint image rather than the core/delta points.

8. It is not clear why synthetic fingerprints are used in this authentication protocol, but not in bio-crypto. After all, attack can happen even during the fingerprint capturing stage. Is there any key being used in the synthetic fingerprint process?

   Response:
   We generate a session key using two different fingerprint image. In order to generate same session key between C and AS, and C and RS, we keep same set of synthetic fingerprint images with AS and RS. We use synthetic fingerprint image because it is stored in remote location. If synthetic fingerprint image is compromised, then it does not compromised security for server as well as for a user. We can easily replace the synthetic fingerprints and no additional security is required to store synthetic fingerprint image. We do not use any key in the synthetic fingerprint process.

9. In chapter 4, you are proposing a technique to generate a bio-code for encryption and the objectives are to have a distinctive code and good similarity in feature vectors. These are also part of the objectives in the bio-crypto in Chapter 3. That is why both Tables 3.3 and 4.4 are identical. Why can't you just use bio-crypto for your use authentication protocol? You should try to shorten Chapter 4 by describing the difference between the two key generations. Do not reproduce materials or tables that are described in the earlier chapters.

   Response:
   To make the Chapter 4 shorten, Table 4.4 has been removed from the Chapter 4. Note that the key generation method as discussed in Chapter 3 and 4 are totally different. Hence we keep these as they are.

10. In Table 5.2, there are two methods cited for fingerprint alignment, consistent region selection and minutiae point's selection. It is not clear how the author integrated these techniques to get the accuracy figures in this table.

    Response:
    Our approach consider the following steps: fingerprint alignment (step-1), consistent region selection (step-2), horizontal segment selection (step-3), and Trellis diagram generation (step-4). We integrate the existing techniques as follows: First we use existing fingerprint alignment mechanism to align the fingerprint image and then follow our proposed steps 2-4. Second, we follow proposed step-1, and then we use existing consistent region and minutiae points selection, the remaining proposed steps 3-4. This way we integrate the existing mechanisms to get the accuracy figures in this table.

11. Instead of designing a new multi-party authentication protocol, did the author try to propose using fingerprints to generate the keys in the existing protocols? What is the major contribution in this chapter, the bioID, the protocol system design or both?

    Response:
    No, we have not tried to propose using fingerprint to generate the keys in the existing protocol. There are two major contributions of this chapter. First, is to generate bioID and second, the new multi-party authentication protocol.

Minor Amendment:

1. Page 59, "which is belongs to" should be rephrased?

   Response:
   We have replaced "which is belongs to" by "belongs to". Please see the third Para of Page 61 in the modified thesis.

2. Page 60, change "is belongs to" to "belongs to"

   Response:
   We have replaced "is belongs to" by "belongs to". Please see the third Para of Page 61 in the modified thesis.

3. How to get 25×10000 years on page 60? Provide reference for this figure.

   Response:
   We thankful to the examiner for suggesting this points. For 5714×10^{16} brute-force trials requires $906303 \approx 90 \times 10^4$ years [W. Stallings, Cryptography and Network Security: Principles and Practice, 5, Ed. Pearson, 2008.].

   We have incorporated the above calculation in the modified this. Please see Last Para of Page 60 in the modified thesis.

4. In Section 3.4.1.3, it sounds like minutiae based technique has problem due to threshold setting in the feature matching. However, the proposed SVM based ranking is also minutiae based. A poor extraction of the minutiae will influence the accuracy of the SVM based ranking. The difference is that one uses feature and the other uses score in the fusion process.

   Response:
   There are following advantages of the SVM based ranking.

   (a) No need to use threshold value. The selected threshold value may not be accurate.

   (b) In SVM based ranking, we have accurate decision boundary. Suppose, T be the threshold value and let the comparison result is close to T but not satisfy the T, then the genuine user will not be authenticated. While in SVM based ranking, if the user is genuine, we SVM ranking gives positive value even if the score (i.e., positive score) is $< T <$ for a genuine user. Hence, SVM based ranking is more accurate than the traditional feature based authentication.

   (c) More secure than the traditional authentication as there is no way to change the threshold value.

5. In 3.4.2, it is suggested that attacker needs to make 2 to the power of 1024 for the 1024 bit key. This is true for the traditional encryption key generation which is random in nature. For the proposed bio-crypto key, core and delta points usually appear in the predicated regions. The histogram pattern of the distance between minutiae and these core or delta points can be observed. Therefore, the bio-crypto key generated is not fully random.

   Response:
   Our key generation algorithm does not depends on the appearance of the core and delta point. Hence, in the predicated region, the existence/non-existence of the core/delta point cannot be used to observe any key pattern. In fact, the proposed key generation

algorithm is least concern about the patten of the distance between minutiae and the core/delta points. It may be noted that we do not use the distance between minutiae and core/delta to generate the key. Hence, the proposed bio-crypto key is fully random.

6. Page 64, "..In our approach.."?

   Response:
   We have removed "In our approach" from the thesis. Please see second para of Page 67 in the modified thesis.

7. In 3.4.4, define TP as True-Positive first

   Response:
   We have defined TP as True-Positive in the modified thesis. Please see Para 2, Page 69 in the modified thesis.

Response to the comments of Indian Examiner (Prof. Phalguni Gupta, IIT Kanpur)

1. Suppose there are two delta points in the fingerprint. One has been captured at the time of registration while other one is seen at the time of authentication. It is not mentioned in the thesis how it will be handled in that case. It needs to be elaborated in the thesis and such cases are to be considered in the experiments.

   Response:
   The same concern has been raised by the external examiner and we have already addressed. (Page 2)

2. It is not necessary that same minutiae points from a fingerprint will be extracted every time. For example, at the time of registration, there are $M1$ minutiae points have been extracted by the extractor. Out of these minutiae points, $A1$ minutiae points are genuine and remaining $M1 - A1$ are false. At the time of authentication, the user gives his fingerprint which contains $N1$ minutiae points. Out of these $N1$ points, $B1$ of them are true minutiae and are from $A1$; $B2$ points are new true minutiae, and $B3$ of them are false minutiae. Among $B3$ points, some of them are found at the time registration. Getting all true minutiae points are itself a challenge. This scenario has to be addressed clearly and elaborately in thesis. Corresponding cases should be analyzed in the experiments.

   Response:
   In biometric, the detection of minutiae points is mostly uncertain. Hence, removal of spurious minutiae points may not be guaranteed to keep original minutiae points. Considering this fact, we process all the captured minutiae points of the captured fingerprint image. In our approach, if some genuine/spurious minutiae are present/absent in the captured fingerprint image during encryption/decryption phase, it would not affect the accuracy in our approach.

   We consider the following cases:

   (a) Case 1: If fingerprint I_1 does not contain any spurious minutiae points, while I_2 contains spurious minutiae points and vice-versa:
   In this situation, the fingerprints I_1 and I_2 will generate the set F_B and F'_B, respectively. Here, F_B contains length ratios and angle differences from genuine minutiae points. But, F'_B contains few length ratio and angle difference which are generated by spurious minutiae points. Therefore, during the comparison of F_B and F'_B, the length ratio and angle difference of F'_B will not be matched with any length ratio and angle differences of F_B. Lets consider this score as S_1 (Fig. 1). Note that S_1 is a score which is calculated using genuine length ratios and angle differences of I_1 and I_2. Also, S_1 will always be created in any fingerprint image.

   (b) Case 2: If fingerprint I_1 does contains core point but I_2 does not contain the core point:
   In this situation, the sets F_C and F'_C will not be generated. Hence, the authentication will be based on the scores S_1 and S_3.

   (c) Case 3: If fingerprint I_1 does contains delta point but I_2 does not contain the delta point:
   In this situation, the sets F_D and F'_D will not be generated. Hence, the authentication will be based on the scores S_1 and S_2.

   (d) Case 4: If fingerprint I_1 and I_2 does not contain the core/delta point:
   In this situation, the authentication will be based on the scores S_1 only.

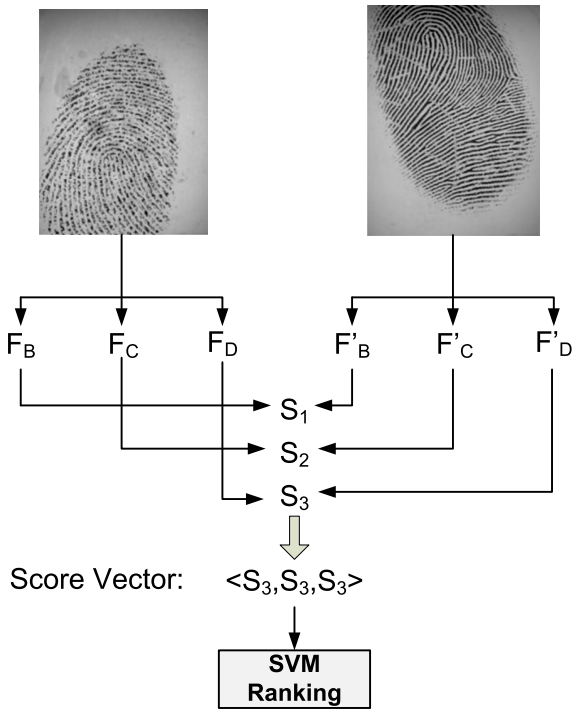

Figure 1: Score vector calculation from different fingerprint images.

This score vector ($< S_1, S_2, S_3 >$) will be passed to our SVM Ranking mechanism. The overall ranking of the score vector will be calculated. If the overall ranking of the score vector is positive then the user is genuine else impostor. Therefore, presence of the false minutiae/core/delta points will be affect the accuracy.

3. There are several typographical errors in the thesis. For example,

   (a) In Page 43, "Later, C is use to verify.." should be written as "Later, C is used to verify...".

   Response:
   The statement has been modified in the revised thesis. Please see Para 1, Page 45 in the modified thesis.

   (b) In Page 44, "Let F_1 represents..." should be "Let F_1 represent.."

   Response:
   The statement has been modified in the revised thesis. Please see Page 46 in the modified thesis.

   (c) In Page 86, 117, it contains similar type of typos

   Response:
   We have carefully checked all the typos of page 86 and 117. Please see the modified thesis.

4. In Page no. 66 it is written "after the block size 75×75 pixels,..." should be written as "after the block size of 75×75,..."

   Response:
   The statement has been modified in the revised thesis. Please see Para 2, Page 69 in the modified thesis.

5. It is not a good practise to refer a section which has not been discussed apriori. For example, In Page 88 [last line of second para] "... is discussed in Section 4.2.6.". This can be avoided by deleting the line itself.

   Response:
   The statement has been removed from the thesis.

6. Use of subscripts at every stage should be seen carefully. They are misleading. Because of this, the examiner faced a lot of problem in understanding the content. In Page 90, lr_{ij} is the length ratio between d_i and d_j. One should define again the meaning of d_i and d_j. But in Equation 4.25 is has written lr_1. What is the relation between lr_1 and lr_{ij}? Same in the case with ad_{ij} and ad_j.

   Response:
   lr_{ij} represents length ratio calculated using i^{th} and j^{th} lines and lr_1 represents first length ratio out of possible length ratios. To avoid the confusion, we modify the Equation 4.25 as follows.

   $$FV_1 = (lr_{ij}, ad_{ij}) \tag{1}$$

   Here, $i, j = 1$ to z_1, $i \neq j$.

   Please see Equation 4.25 in the modified thesis.

7. In References, there are some papers without issue no. For example, [53], [57].

   Response:
   The references have been modified in the revised thesis.