SECURE BIOMETRICS

By

QINGHAI GAO

A Dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2008

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the

Graduate Faculty in Computer Science in satisfaction of the

dissertation requirements for the degree of Doctor of Philosophy.

Professor Michael Anshel

Date — Chair of Examining Committee

Professor Theodore Brown

Date — Executive Officer

Professor Candido Cabo

Professor Ping Ji

Professor Xiangdong Li

Professor Li-Chiou Chen

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

SECURE BIOMETRICS

By

Qinghai Gao

Adviser: Professor Michael Anshel

Biometric systems attempt to solve a matching problem through live measurements of human body features. One main barrier that prevents the widespread application of biometrics is the concern about the security and privacy of biometric information. To address this concern biometrics needs to be protected with cryptography. However, the specific problems with biometrics, namely number limitation, non-secrecy, non-reproducibility and non-cancelability, make it a challenge to secure biometrics effectively with existing cryptographic algorithms, especially on how to match two biometric templates in encrypted or hashed formats.

Inspired by the Central Dogma of Biology, we in this thesis develop an Artificial Intronization Method (AIM) to approach the problems of biometrics.

AIM is a method of inserting introns into an exon sequence to obtain ciphertext. Three methods are proposed to introduce introns into plaintext: Pseudo Random Number Generator, Integer Sequence and Geometric Key. The main advantage in using AIM is to prevent error propagation. However, one disadvantage of AIM is that security may require a large Message Expansion Rate. Therefore, three methods are proposed to control message expansion: Intron Compression, Intron Removal, and Exon Elimination.

With AIM, the number limitation, non-secrecy and non-cancelability problems can be solved by applying different intronization keys and different symmetric-key based intron sets; in theory, AIM could achieve zero-error propagation – a solution to the non-reproducible problem.

We believe AIM can be an effective hashing mechanism for protecting fuzzy biometrics. Our testing results support this belief. With this finding, AIM can also be used as a preprocessing step for other cryptographic algorithms to enhance security.

AIM, due to the intentional suppression of the diffusion property in favor of zero-error propagation, is vulnerable to Known-plaintext Attack. Thus, it has its limitation as a stand-alone cipher.

ACKNOWLEDGEMENTS

Most of all, I thank my advisor, Professor Michael Anshel, who helped me select the topic that fits my background and interests, and whose knowledge, guidance and encouragement helped me carry out the research and complete this thesis.

I would also like to recognize the following professors for their invaluable help in aiding me to create and then finish my thesis.

· Professor Ted Brown who admitted me into the program and helped me during difficult times

· Professor Candido Cabo, Professor Li-Chiou Chen, Professor Xiangdong Li and Professor Ping Ji for serving in my committee and giving me suggestions on my research

· Professor Lin Leung who gave me endless help on many aspects

· Professor Kent Boklan for his interesting lectures on cryptanalysis

· Professor Bud Mishra for introducing bioinformatics to me

· Dr. Craig Watson for sending me the NIST fingerprint software

In addition, I thank Xiaowen Zhang, Damike Kahanda, and Ke Tang who helped me with my studies.

Also I thank Judy Waldman for helping me with the format of my thesis.

At last, I would like to thank my wife Cheng Zhang and my son Yuan Gao for their patience and support.

To my parents

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1 Problems with biometrics

Abstract: Biometrics has four main problems, number limitation, non–secrecy, non-reproducibility and non-cancelability, which make it necessary and difficult to protect.

1.1 Introduction

Passwords are the most common form of authenticating users. Password-based online authentication works with two stages. During registration, a user selects a password and then the hash of the password will be saved in a server database. During authentication, the newly input password will be hashed locally, and the new hash will be sent over the Internet to the server for comparison with the saved hash. If the two hashes match, authentication is successful, otherwise unsuccessful. In general, a unique username is required for each account. The uniqueness of username guarantees authentication is a 1-to-1 verification process.

The advantage of password-based authentication is that passwords are short text strings without containing any non-deterministic bit, thus can be memorized and hashed in real-time with one-way hash function, like SHA-1.

However, there are a few problems with this scheme. As internet access becomes an essential activity of our daily lives, we have too many passwords to remember. Non-

repudiation is another problem because password can be transferable easily. Therefore, people have been looking for better ways for high security requirement.

One proposed solution to these problems is to use biometric authentication because biometrics are physically with us and can't be transferred easily. In fact, a user of biometrics does not know or may not need to know the details of biometrics. All he needs to remember is which biometrics he uses: whether it is a particular fingerprint, iris, etc.

Like password-based systems, biometric identification works in two stages: enrollment and verification. At both stages, a raw image is obtained with some instruments by measuring a live biometric. The raw image is then used for feature generation. Features extracted are often transformed into a template, which contains less data than the raw image, to facilitate storage and matching processes. The template and the raw image may be stored in a centralized database or distributed on a smart card.

Due to the statistical nature of the acquisition and matching stages, biometric systems are never 100% accurate. Two types of errors are defined for the inaccuracy: false match rate and false non-match rate. These errors vary from one biometric technology to another and also depend on the threshold, which is set based on the security requirements.

The main objectives of biometric recognition are user convenience and better security. We believe that wider applications of biometric technologies are inevitable and necessary as our society demands higher security.

However, biometric applications have raised a series of issues [50-56], which prevent its wide acceptance. Among them **the security and privacy issues of biometric information themselves are considered more important than other issues**. Once in a while, we hear news that databases were hacked and personal data was stolen.

Even though it has been realized in the biometric research community that biometric information needs to be protected with cryptography [1-4], the mainstream research in biometrics has focused on image processing and pattern recognition, trying to improve the accuracy of biometric systems without sacrificing too much security. Research specifically concentrated on protecting biometrics is limited.

1.2 Problems with biometrics

There are four main problems with biometrics: number limitation, non-secrecy, non-reproducibility, and non-cancelability[7].

Number limitation (NL)

We only have a limited number of directly available biometrics. In literature, multi-biometrics [38-47] and partial biometrics [6] [49] [53] are two possible solutions.

Non-secrecy (NS)

Some biometrics can be easily stolen, for example, faces can be photographed. However, biometric raw information collected without the cooperation of a person is often of poor quality. Another layer of protection is that biometric template extraction algorithms are often proprietary, which helps to maintain the secrecy of biometrics. The non-reproducibility of biometrics also helps protect its secrecy.

Considering a large biometric database with millions of users and the biometric information is NOT transparently linked with his or her non-biometric information, we may reasonably assume that biometrics is kept secretly.

Proposed solutions for the non-secrecy problem include multimode biometrics [38-47], partial biometrics [6] [49], anonymous biometrics [5] [53], and combining biometrics with password [4].

Non-reproducibility (NR)

There is inherent fuzziness with biometric measurements due to changes in physical and environmental conditions. This is the most difficult problem for protecting biometrics.

We summarized the common methods used in the image processing and pattern recognition community to approach the fuzzy measurement problem, without referring to any particular paper.

Averaging/Training

During registration, a number of biometric images with some variations are obtained, transformed, and then averaged to get a generic representation of the biometrics. During authentication, the same biometrics will be measured and compared with the mean representation.

Quantization/Tessellation

During registration, a biometric image will be quantized into a number of small units. The information of interest located inside each unit will be assumed to come from the center of the unit.

Majority voting

For an odd number of measurements of a biometric, each will be quantized and binarized into a fixed-length string. For every position of the set of binary strings, majority voting will be used.

Error correcting

Based on the characteristics of the biometric measurements, some redundant information will be added into the final representation, which will be used to correct the inconsistent bits during authentication.

In spite of all these efforts, the measurements of biometrics will never be 100% accurate. The best we can do to approach the non-reproducible problem on transforming a given biometric template for security purpose is to prevent the increase of false match rates due to the cryptographic transformation.

Non-cancelability (NC)

Biometrics is part of us. If a biometrics is compromised, it can't be canceled easily. Recently this becomes a hot research topic in biometric community. Quite a few methods have been proposed in literature [7-35] [48]. However, in essence all these proposals involve combining biometric information with a non-biometric key.

1.3 Summary

Biometrics refers to the recognition of an individual based on behavioral and physiological characteristics, which include face, fingerprint, iris, retina, signature, speech, facial/hand thermogram, hand/finger geometry, hand vein, ear, gait, palmprint, keystroke, DNA, etc.

Security and privacy concerns make it necessary to apply cryptography for the protection of biometric information. To secure biometric information, we need to approach the four specific problems with biometrics, namely, number limitation, non–secrecy, non-reproducibility and non-cancelability. In literature methods have been proposed to approach these problems. However, research in this area is still at its early stage. An effective solution has yet to be seen. This thesis makes an effort to approach these problems.

Chapter 2 Artificial Intronization Method

Abstract: Artificial Intronization Method (AIM) is a method of inserting introns into an exon sequence to obtain ciphertext. Three methods are proposed to introduce introns into plaintext: PRNG, Integer Sequence and Geometric Key. Exon Elimination is proposed to combine exons with introns. The security of AIM is analyzed.

2.1 Introduction

According to Shannon [58], a strong cipher should have good diffusion property and confusion property. To achieve them some techniques would be applied during encryption such as substitution, permutation/transposition, combination, fractionation, etc. Existing private-key algorithms (e.g., AES), public-key algorithms (e.g., RSA), and cryptographic hash functions (e.g., SHA-1 and MD5) have good diffusion and confusion properties. However, they may not work well for securing biometrics due to the fuzzy measurements of biometrics if we match biometric templates in encrypted or hashed formats.

In Biology, the genes of eukaryotes have introns that separate exons [63]. A pre-mRNA transcript is made directly from a gene, the introns are sliced out, and exons are joined sequentially to form mRNA, which leaves the nucleus. Translation of mRNA into protein occurs in the cytoplasm.

In a eukaryotic cell, only less than 10% of the entire DNA sequence is directly used for protein coding. That is to say, the majorities of DNA are introns, which were once called junk DNA. Modern biologists believe that introns play important roles. However, finding the exact roles of introns are ongoing research problems. From the

security point of view, we believe that the protection due to the existence of introns or non-coding regions in DNA help organisms to survive. It is generally agreed that more advanced organisms have more non-coding regions in their DNA.

Inspired by the intron removal process, we introduce an Artificial Intronization Method (AIM) as a cryptographic technique, which adds introns into plaintext to obtain ciphertext.

2.2 Artificial Intronization

Different methods can be used to put introns into a plaintext message. In this thesis we introduce three methods: **Pseudo Random Number Generator** (PRNG), **Integer Sequence** and **Geometric Object**.

2.2.1 PRNG

For a given seed, a PRNG can produce a binary string, which contains roughly equal number of 0s and 1s. Then we can let all the 0s be introns and 1s be exons, or vice versa.

Table 2.1 gives an example. For a biometric template, e.g., ATTGCGGATC, we can intronize it as given in the 3rd row of Table 2.1. Letter X can be A, T, G, or C.

In Table 2.1, the 4th row is obtained with **Binary Intronization**: the 1s and the 0s of the 2nd row are further divided into introns and exons, respectively.

Table 2.1 Artificial Intronization with Pseudo Random Number Generator

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	0	1	0	1	1	0	0	0	1	0	1	1	1	1	0	0	1	1	0
X	X	A	X	T	T	X	X	X	G	X	C	G	G	A	X	X	T	C	X
0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0	1	0	0	1
X	A	X	T	T	X	G	X	X	C	X	G	G	X	A	X	T	X	X	C

2.2.2 Integer Sequences

The On-Line Encyclopedia of Integer Sequences [60] contains hundreds of thousands of integer sequences. We can select an entire sequence or one section of a sequence, modify it and then use it to specify the positions of introns and exons.

Example 1(A054646)-Hailstone sequence

- If n is even, divide it by 2 to give n' = n/2.
- If n is odd, multiply it by 3 and add 1 to give n' = 3n + 1.

n = 11 gives the sequence

11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, ...

n=231 gives the sequence

231, 694, 347, 1042, 521, 1564, 782, 391, 1174, 587, 1762, 881, 2644, 1322, 661, 1984, 992, 496, 248, 124, 62, 31, 94, 47, 142, 71, 214, 107, 322, 161, 484, 242, 121, 364, 182, 91, 274, 137, 412, 206, 103, 310, 155, 466, 233, 700, 350, 175, 526, 263, …

For the sequence of n=231 we can modify the sequence (e.g., mod N, N is the length of ciphertext) or select a subsequence.

Example 2 (A005150): Look and Say sequence

1, 11, 21, 1211, 111221, 312211, 13112221, 1113213211, 31131211131221, 13211311123113112211, 11131221133112132113212221, 3113112221232112111312211312113211, 1321132132111213122112311311222131111222113111221131221

Note, instead of using the entire sequence, we can just pick one of the numbers, do some modification, change it into binary, and use the resultant sequence.

Example 3: use π for intron position

π=3.1415926535 8979323846 2643383279 **5028841971** 6939937510 [62]

Pick the highlighted group of digits for intronization,

Group I=0101 0000 0010 1000 1000 0100 0001 1001 0111 0001

Also we can do transformations as the following,

1/sqrt(π)=0.5641895835 4775628694 **8079451560** …

Then pick the highlighted group and translates it digit-by-digit into binary,

Group II=1000 0000 0111 1001 0100 0101 0001 0101 0110 0000

Take the complement of the XOR of the Group I and II,

Group III=0010 1111 1010 1110 0011 1110 1110 0011 1110 1110

Group III will be used to specify intron and exon positions.

According to Sloane [61], about 10,000 new sequences are added into the database annually. We anticipate that it might be difficult to carry out dictionary attacks because of the large number of sequences and the symmetric-key based modification.

2.2.3 Geometric Key

Geometric objects can be selected or designed to act as the key(s) for the intronization technique, as shown in the following examples.

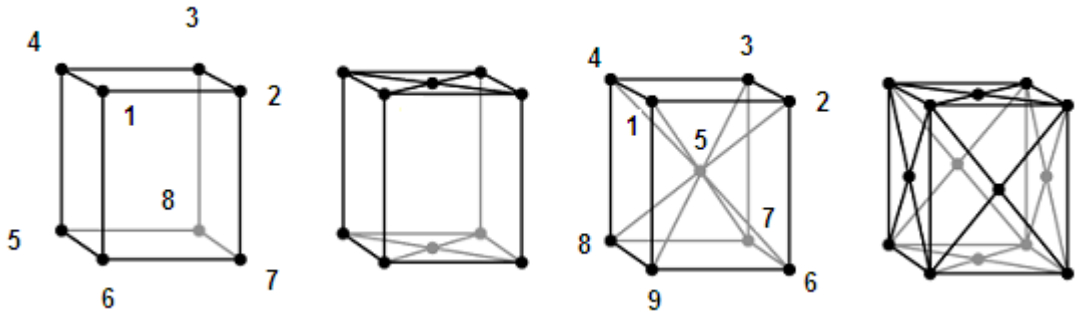

Figure 2.1 Some geometric keys for intronization

Example 1: use Cube1 to wrap a sequence without intronization.

By wrapping the sequence along the vertices of Cube 1 and then reading out vertically, we obtain the results given in Table 2.2.

Table 2.2 Cube1, 5 rounds, counter-clockwise (Top view) wrapping*

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
II	1	6	11	16	2	7	12	13	3	8	9	14	4	5	10	15
III	1	7	9	15	6	12	14	4	11	13	3	5	16	2	8	10
IV	1	14	11	8	12	5	2	15	3	16	9	6	10	7	4	13
V	1	7	9	15	6	12	14	4	11	13	3	5	16	2	8	10
VI	1	14	11	8	12	5	2	15	3	16	9	6	10	7	4	13

*Refer to Figure 2.2 for Row I, II and III

Unlike simple transposition with fixed period, the rotation and the transposition happen concurrently because of the choice of the cube and the way of wrapping. For this example, the block size is 16.

More rounds can be applied. However, the sequence repeats every other round. Therefore, the simple geometric wrapping without intronization is not secure. Note that we don't count the 1st row of each table as a round. The 1st round starts at the 2nd row.

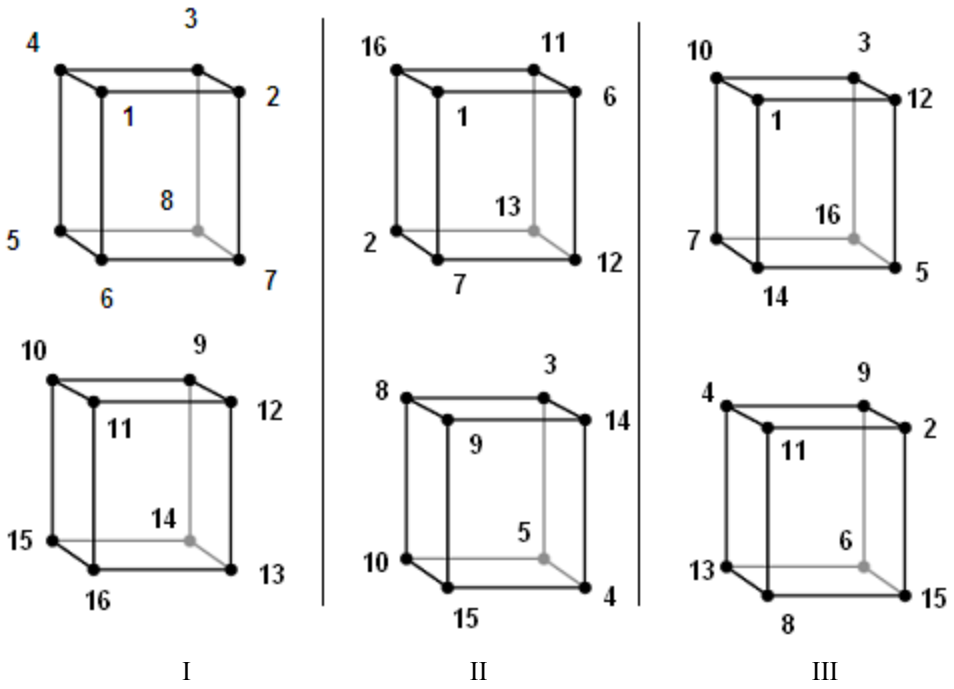

Figure 2.2 Row I, II and III for Table 2.2

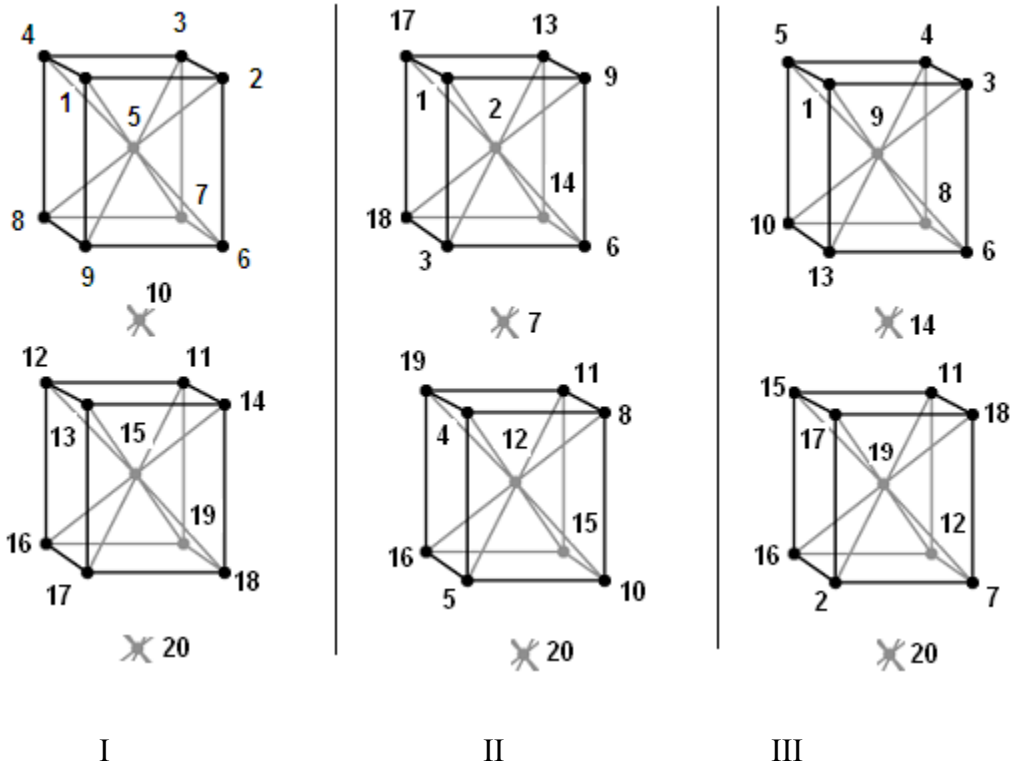

Figure 2.3 Row I, II, and III for Table 2.3

Example 2: use Cube 3 to wrap a sequence with intronization.

The sequence is wrapped around the Cube 3 counterclockwise (Top view), and then read out vertically to form the new sequence. Since the center position of the cube is appointed as intron position, 5, 10, 15 and 20 are all introns (gray-highlighted). With four rounds there are 7 exons and 13 introns, as given in Table 2.3.

Table 2.3 Cube 3, counter-clockwise (Top view), 4 rounds*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	9	13	17	2	6	14	18	3	7	11	19	4	8	12	16	5	10	15	20
1	3	4	5	9	6	8	10	13	14	11	15	17	18	19	16	2	7	12	20
1	4	5	9	13	6	10	14	17	18	11	19	2	7	12	16	3	8	15	20
1	9	13	17	2	6	18	7	3	8	11	15	4	10	19	16	5	14	12	20

*Refer to Figure 2.3 for Row I, II and III

In the 5th row of Table 2.3, five positions, 1, 6, 11, 16, and 20, never change values due to the one directional wrapping (The period 5 reveals some information about the geometric object).

If we use a more sophisticated wrapping, e.g., wrapping counterclockwise and then clockwise alternatively, the results are given in Table 2.4. With four rounds there are 5 exons and 15 introns.

Table 2.4 Cube 3, counterclockwise and then clockwise (Top view), 4 rounds

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	7	11	17	2	6	12	16	3	9	13	19	4	8	14	18	5	10	15	20
1	12	13	5	7	6	19	18	11	3	4	15	17	16	8	10	2	9	14	20
1	15	17	7	19	6	14	9	4	13	5	8	2	10	18	3	12	11	16	20
1	18	12	14	16	6	10	13	7	2	19	9	15	13	11	17	8	5	3	20

With this way of wrapping, two exons (1 and 6) and one intron (20) remain fixed in their positions. One way to overcome this problem is to ***rotate the sequence*** before wrapping it around the Cube.

In Table 2.5, the 3rd row is the rotation sequence, not one of the rounds. The 4th row is the rotational result of the 2nd row, and the 6th row is the rotational result of the 5th row.

From these results we can see that one disadvantage of the algorithm is that more rounds means less positions can be used for plaintext if all other factors are the same. For example, in the 7th row there are only four exons (7, 18, 11, and 13). All others are introns. Therefore, the M**essage Expansion Rate** (MER) equals to 20/4=5, which means on average every symbol has 20 percent of a chance to be an exon and 80 percent to be an intron. In the 5th row there are eight exons (13, 1, 18, 12, 11, 16, 19 and 7), thus MER=20/8=2.25.

Table 2.5 Cube 3, rotating by 4 positions, 3 rounds (IV and VI are rotations)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	7	11	17	2	6	12	16	3	9	13	19	4	8	14	18	5	10	15	20
5	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*1*	*2*	*3*	*4*
2	6	12	16	3	9	13	19	4	8	14	18	5	10	15	20	1	7	11	17
2	13	14	1	6	9	18	20	12	4	5	11	16	19	10	7	3	8	15	17
6	9	18	20	12	4	5	11	16	19	10	7	3	8	15	17	2	13	14	1
6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1

A few methods can be used to control the MER, including **Exon Elimination** (section 2.3), **Intron Compression** and **Intron Removal** (Refer to Chapter 3 for more information).

One advantage of the AIM is that the geometric object can be designed arbitrarily. And for a selected geometric object we can select the intron positions arbitrarily. For example, if we choose vertex 1, 7 and 17 as intron positions and follow the same steps as those of Table 2.5, we will obtain a different sequence as given in Table 2.6.

Table 2.6 Cube 3, rotating by 4 positions/counterclockwise, 5 rounds

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	7	11	17	2	6	12	16	3	9	13	19	4	8	14	18	5	10	15	20
5	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*1*	*2*	*3*	*4*
2	6	12	16	3	9	13	19	4	8	14	18	5	10	15	20	1	7	11	17
2	13	14	1	6	9	18	20	12	4	5	11	16	19	10	7	3	8	15	17
6	9	18	20	12	4	5	11	16	19	10	7	3	8	15	17	2	13	14	1
6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1

In the 7th row of Table 2.6 there are 10 exons and 10 introns. Therefore, on average, every symbol has a 50 percent of a chance of being an exon.

For a given biometric template, e.g., ATTGCGGATC, if we transform it with AIM and using same key(s), many different ciphertexts can be generated by simply changing the introns, as shown in Table 2.7 (The 2nd row is the 7th row of Table 2.6). Therefore, **cancelable biometrics** can be achieved.

Table 2.7 Many ciphertexts for one biometric template (X=A, T, C or G)

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1
III	X	X	A	X	T	T	X	X	X	G	X	C	G	G	A	X	X	T	C	X

The main advantage of the AIM is that there is no error propagation, which effectively solves the **non-reproducible** problem of biometric measurements.

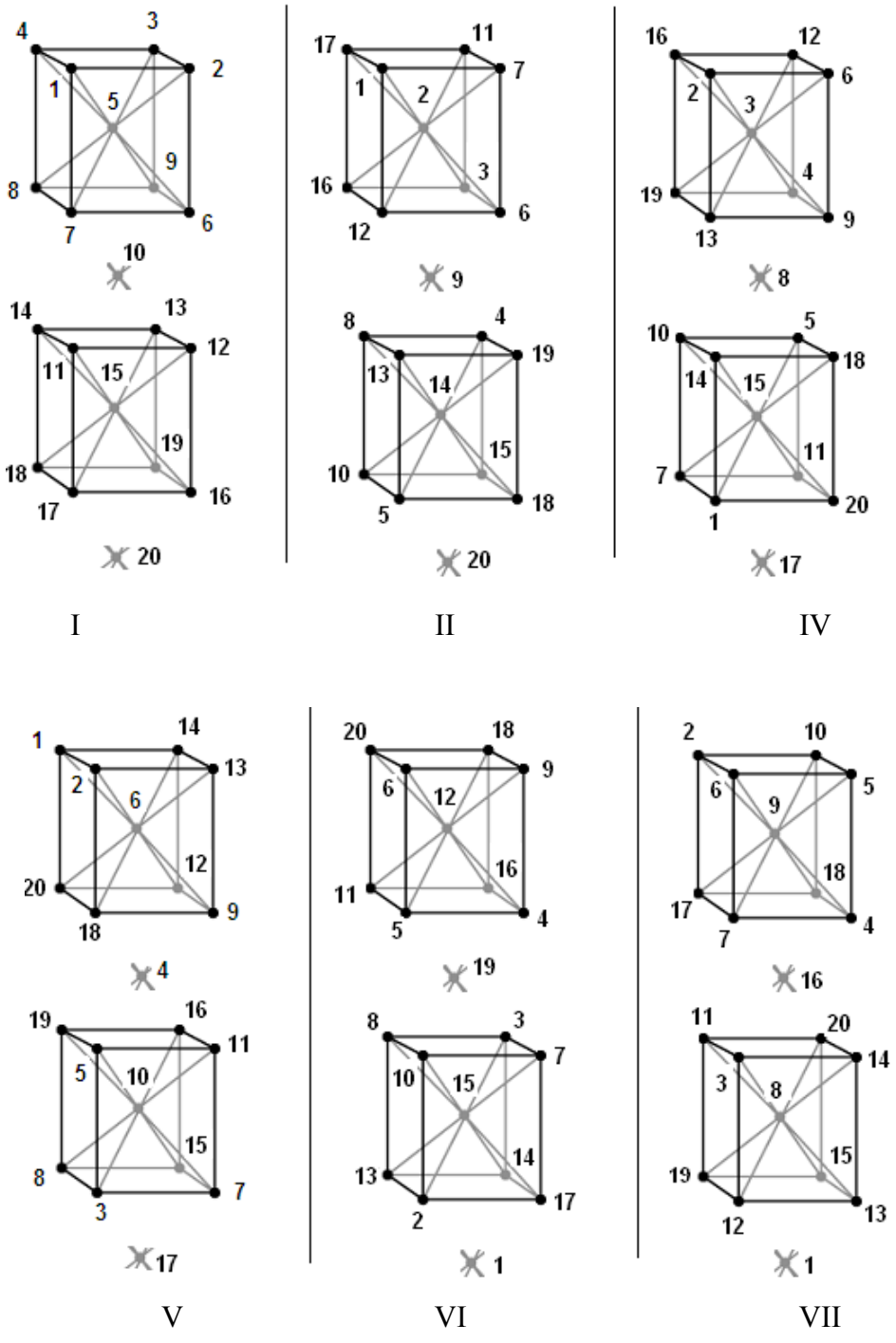

Figure 2.4 A few rows for Table 2.5

As mentioned above, different geometric objects can be chosen or designed to increase the security of the scheme.

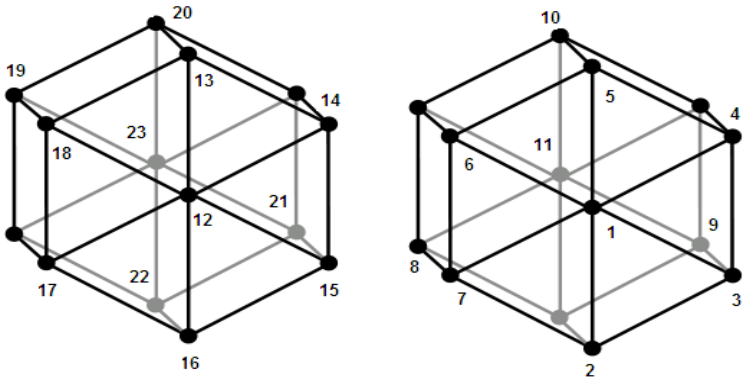

Figure 2.5 Hexagonal Key for Artificial Intronization

Fr example, if we use a hexagon to wrap a sequence according to Figure 2.5, the following sequence will be generated with one round:

1, 11, 12, 23, 2, 16, 22, 3, 9, 15, 21, 4, 14, 5, 10, 13, 20, 6, 18, 19, 7, 8, 17

Generally, the encryption key may consist of the following information:

(1) Geometric object and intron positions

(2) Wrapping direction, clockwise or counterclockwise

(3) Number of rounds

(4) Other operations, like shifting

2.2.4 Combining geometric object with pseudo random sequence

One problem with our wrapping in Section 2.2.3 is that there may not be enough randomness in the output. To approach this problem, we can use pseudo random sequence to guide the wrapping.

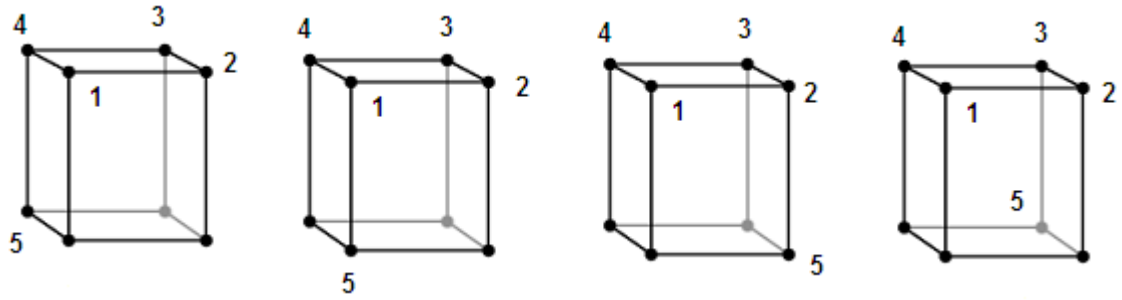

Figure 2.6 Wrapping guided by pseudo random sequence

Assume we wrap a sequence vertex-by-vertex with Cube 1 (Refer to Figure 2.6). After we reach vertex 4, there will be four different choices for the vertex 5. Once we choose the position for vertex 5, wrap the other three vertices on the same square until we are about to wrap on the next square. In this example, there will be 16(n-1) different ways to wrap it for a sequence with a length of 4n.

With this method we need another sequence to specify which vertex to select at each level. And this level-wised vertex selection sequence can be generated with PRNG (section 2.2.1), or by using a modified sequence from the encyclopedia (section 2.2.2).

2.3 Exon Elimination

So far, the ciphertexts obtained in Section 2.2 are mixtures of introns and exons. Since exons contain fuzzy bits, arithmetic operations between exons will unavoidably propagate errors. Therefore, we need to avoid combining two exons.

Unlike exons, introns do not contain any fuzzy bits. Therefore, XORing an exon with an intron will not introduce new error bits. Based on this analysis, we introduced a

technique of dissipating exon into intron, and called it **Exon Elimination**. Table 2.8 gives such an example (A=10, C=00, G=11, T=01, operation XOR)

Table 2.8 Exon Elimination - Dissipating exon into intron

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	A	C	G	A	T	A	G	G	C	T	T	G	A	G	T	C	C	A	G	C
III	T	T		C			A	C	A		A					T	C	A	G	C
IV	T	T	C	A	C	A	A	T	C	A	G	C								

Followed are the calculations for the 3rd row:

1⊕3=A⊕G=T, 2⊕5=C⊕T=T, 4⊕6=A⊕A=C, 7⊕10=G⊕T=A, 8⊕12=G⊕G=C, 9⊕13=C⊕A=A, 11⊕14=T⊕G=A, 15⊕16=T⊕C=T

The 4th row is obtained by right shifting.

With Exon Elimination, an exon set can be completely hidden in an intron set. In addition, Exon Elimination can be utilized to control message expansion.

2.4 Evaluation of the Artificial Intronization Method

Shannon [58] listed five criteria to estimate the value of a proposed secrecy system:

- Amount of secrecy (Less is better)
- Size of key (Smaller is better)
- Complexity of enciphering and deciphering operations (Simpler is better)
- Propagation of errors (Less is better)
- Expansion of message (Less is better)

As Shannon [58] pointed out, it is very difficult to achieve good results for all five criteria, but it is not difficult to achieve four of them. Table 2.9 gives a rough comparison of different cryptographic techniques.

Table 2.9 Comparison of AIM with other cryptographic techniques

Cipher	**AES**	**RSA**	**OTP**	**Vigenere**	**Intronization**		
					PRNG	Sequence	Geometry
Amount of secrecy	key	Private key	Same length as message	Key	Seed	Sequence name	Object, Intron positions
Size of key	128, 192, 256	512, 1024, 2048	Variable	Variable	Variable	Variable	Variable
Operation complexity	**	**	*	*	**	*	**
Propagation of errors	Yes	Yes	No	No	No	No	No
Expansion of message	No	No	No	No	Yes	Yes	Yes

From Table 2.9, we can see that the main advantage of the AIM is its zero-error propagation. Its main disadvantage is message expansion.

However, Shannon [58] proposed the five criteria about 60 years ago, when the computer had very limited processing power, very limited memory and storage. Since modern computers have far superior processing capability, message expansion should be applied for information security purpose.

We evaluate the security of AIM with the common methods of cryptanalysis, as given in Table 2.10.

In essence the AIM is a symmetric-key based method, therefore it suffers from CPA, CCA, ACPA and ACCA.

Table 2.10 Security of AIM under cryptanalysis

Ciphertext-only Attack (COA)	Secure
Known-plaintext Attack(KPA)	Relatively secure
Chosen-plaintext Attack(CPA)	Vulnerable
Chosen-ciphertext Attack(CCA)	Vulnerable
Adaptive chosen-plaintext Attack(ACPA)	Vulnerable
Adaptive chosen-ciphertext Attack(ACCA)	Vulnerable

However, it may not be very easy to launch KPA for the following two reasons. First, exact plaintext of a biometric template is not easy to obtain by stealing biometric image. Second, Exon Elimination could make ciphertext secure.

Our effort has been focused on developing AIM into a secure method against COA without using Shannon's diffusion property.

AIM can also be applied to enhance the security of substitution ciphers against the common frequency analysis attack.

As Shamir stated in his 2002 Turing Award lecture [59], the three laws of security are:

- Absolutely secure systems do not exist
- Cryptography is typically bypassed, not penetrated
- To halve your vulnerability, you have to double your expenditure

Intronization, an information securing technique developed by nature in billions of years, should have been used more widely.

Chapter 3 Solutions

Abstract: With the Artificial Intronization Method we propose solutions to the four problems of biometrics.

3.1 Number limitation problem

Every intronized biometric template consists of two sets. One set is the exon set derived from a raw biometric image, another set is the intron set generated pseudo randomly. Therefore, one way to generate multiple intronized biometric templates from one original biometric template is to use different intron sets.

Table 3.1 gives an example. Here we use the protein alphabet. Since every X can have 20 different choices, we can have large number of different intronized templates.

Table 3.1 One original template corresponds with many intronized templates*

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1
III	X	X	R	X	D	R	X	X	X	G	X	G	B	G	I	X	X			X

*2nd row is the 7th row of Table 2.6.

For security, both intron set and the positions of introns should be changed. Exon Elimination can also be applied.

3.2 Non-secrecy problem

If an attacker steals a biometric image, it is still not easy for him to generate the intronized template without knowing the introns and keys.

3.3 Non-reproducible problem

Except for DNA, RNA and protein, the measurements for other biometrics are inherently noisy. A practical solution to this problem would be preventing error propagation in the cryptographic transformation of biometric templates. The AIM is designed to achieve this goal.

The intronization process should not change matching score if we just apply it as an encryption method, which means introns will be removed before matching can be done. Due to the removal every intron in ciphertext can randomly take on any value from a chosen alphabet. Table 3.2 shows an example.

Table 3.2 Randomized intronization for encryption

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1
III	A	P	R	V	D	R	L	I	S	G	T	G	B	G	I	N	C	A	G	D
IV	A	P	R	F	K	R	H	I	R	G	K	G	B	G	V	M	L	A	G	B

If we use the same set of introns, the effects of intronization can be illustrated with Table 3.3.

Table 3.3 Insertion of same set of introns for hashing

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1
III	A	P	R	V	**D**	R	L	I	S	G	T	G	B	G	**I**	N	C	M	Q	D
IV	A	P	R	V	**K**	R	L	I	S	G	T	G	B	G	**V**	N	C	M	Q	D

For the 3rd and 4th rows in Table 3.3, the exon-only matching score is 6/8=75%, while the exon-plus-intron matching score is 18/20=90%. Therefore, adding same introns will increase the false match rate.

To prevent an attacker from regenerating the raw biometric template from an intronized one, i.e., to achieve the hard-to-invert property, we proposed two methods. One is to use **Intron Compression,** as shown in Table 3.4. Note that the 1st and 2nd row are copied from Table 2.5.

In Table 3.4, we apply Intron Compression to 3rd row, for example, by XORing continuous introns, to generate the 4th row. The 5th row is obtained by packing to the left. The non-invertibility is achieved because it is difficult to map from the 4th or the 5th row back to the 3rd row.

Table 3.4 Illustration of Intron Compression*

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1
III	1	0	1	1	0	0	1	1	0	1	1	1	0	1	1	1	0	0	1	1
IV				1			1	1	0			1		1	1	1		0		
V	1	1	1	0	1	1	1	1	0											

* Ciphertexts are binary and introns are gray in color.

The second method is to use **Intron Removal,** as shown in Table 3.5. Note that the 1st and 2nd rows are copied from Table 2.6.

Table 3.5 Illustration of Intron Removal*

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
II	6	5	10	2	9	4	7	17	18	16	3	14	20	11	8	13	12	19	15	1
III	1	0	1	1	1	0	1	0	0	0	0	0	1	0	1	1	1	0	1	0
IV	1	0	1	1	1	0				0	0	0	1	0	1	1	1	0	1	0
V				1	0	1	1	1	0	0	0	0	1	0	1	1	1	0	1	0

* Ciphertexts are binary and introns are gray in color.

In Table 3.5, the 3rd row is the original intronized biometric template. Assume we choose to remove the introns whose lengths are equal to or greater than a number, e.g., 3. The result is given in the 4th row. Right shifting gives the 5th row.

Both Intron Compression and Intron Removal can be applied to reduce MER.

3.4 Non-cancelable problem

If a database containing intronized biometric templates is compromised, we can replace these templates by changing introns and keys. Therefore, the cancelability of biometrics can be achieved with the AIM.

Chapter 4 Experimental Results

Abstract: The AIM, as a hashing mechanism, is tested for two different situations: verification (1:1) and identification (1:N). The results prove its effectiveness for protecting fingerprint minutiae templates.

4.1 Verification (1:1)

Six fingerprints from FVC2004 database DB1 [71] are randomly selected. FP34_2a has 25 minutiae. Five minutiae of FP34_2a are significantly modified to obtain FP34_2b. All the minutiae of FP34_2b are slightly modified to obtain FP34_2c. The matching scores for the 8 fingerprints are given in Table 4.1.

Table 4.1 Matching scores for fingerprints from DB1

FP	**1_1**	**8_2**	**34_2a**	**34_2b**	**34_2c**	**65_3**	**97_2**	**105_3**
1_1	499	5	0	0	3	6	0	3
8_2		486	3	3	3	6	5	3
34_2a			*103*	*62*	54	6	3	0
34_2b				104	93	4	3	0
34_2c					104	3	3	3
65_3						499	3	12
97_2							136	5
105_3								219

The lower left half of Table 4.1 is left empty due to the symmetry of matching scores. To test how the AIM changes matching scores we considered six different situations as given below.

4.1.1 With or without introns

FP34_2a, which has 25 original minutiae, is matched against itself inserted up to 50 introns from FP1_1. The results are given in Figure 4.1.

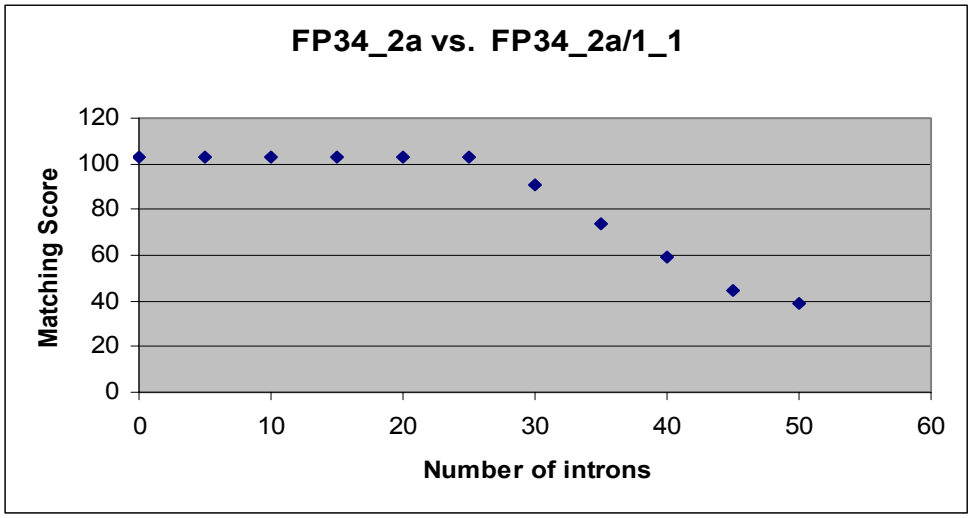

Figure 4.1

From Figure 4.1, we can see that adding up to 25 introns (MER=2) has nearly no effect on the matching scores. However, adding 50 introns (MER=3) reduces the matching scores to ~ 40, the threshold set by NIST [72].

For FP34_2a we insert up to 80 introns from FP105_3 and 65_3, and match them against the original. The results are given in Figure 4.2.

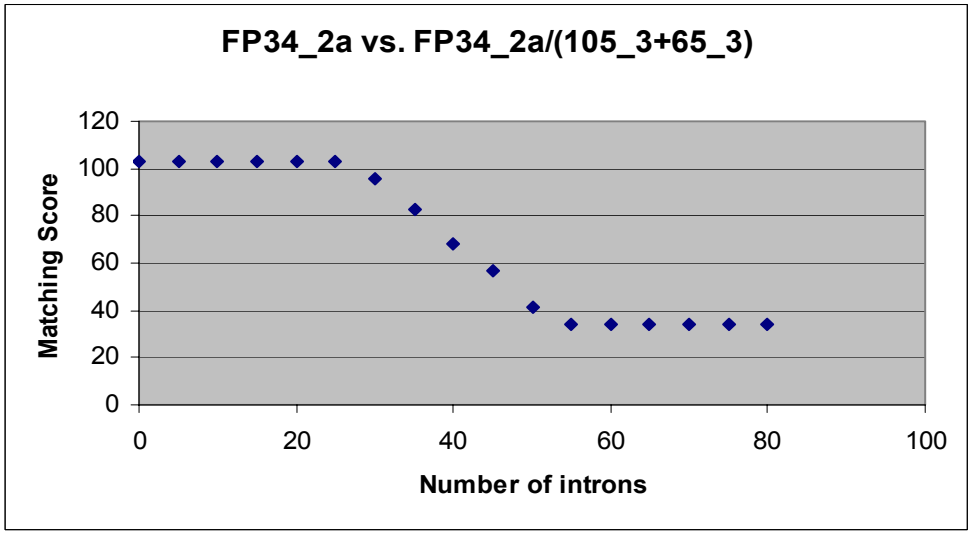

Figure 4.2

FP34_2b, which has 25 original minutiae, is matched against itself inserted up to 50 introns from FP1_1. Figure 4.3 shows the results.

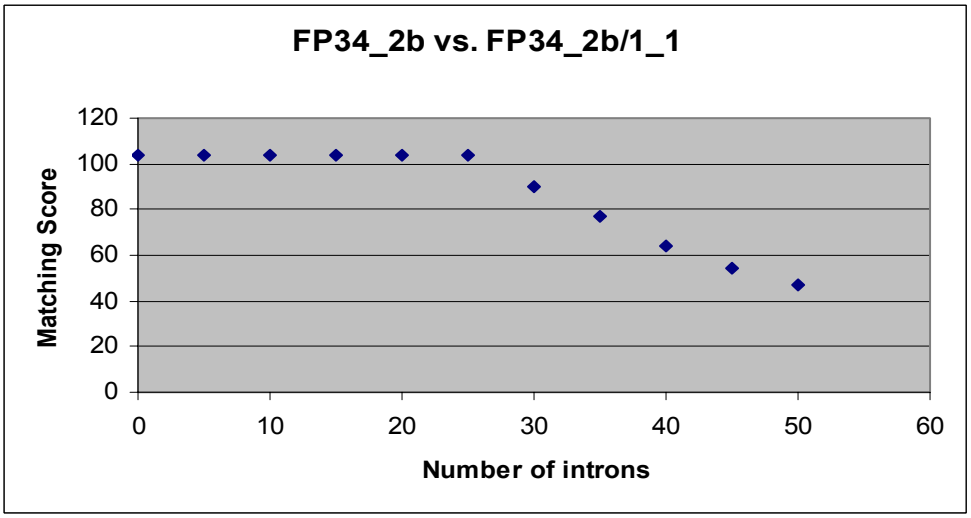

Figure 4.3

FP97_2, which has 25 original minutiae, is matched against itself inserted up to 75 introns from FP1_1 and 8_2. Figure 4.4 shows the results.

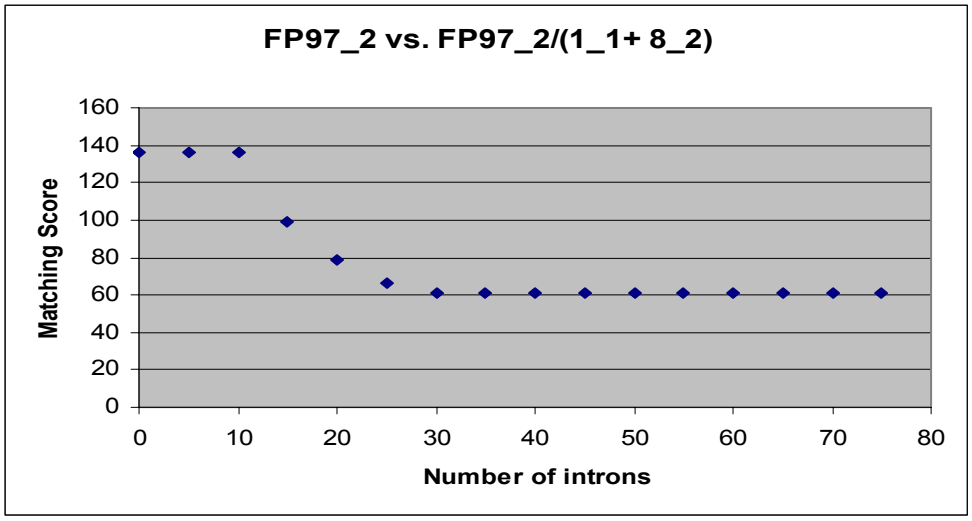

Figure 4.4

The results from Figure 4.1 to 4.4 prove that the AIM can be used as a hashing mechanism to protect fingerprints because the fingerprint templates inserted with certain number of introns can still match the original.

4.1.2 Same fingerprint with different sets of introns

For FP34_2a, we inserted the same numbers but two different sets of introns, then match them against each other. The results are given in Figure 4.5.

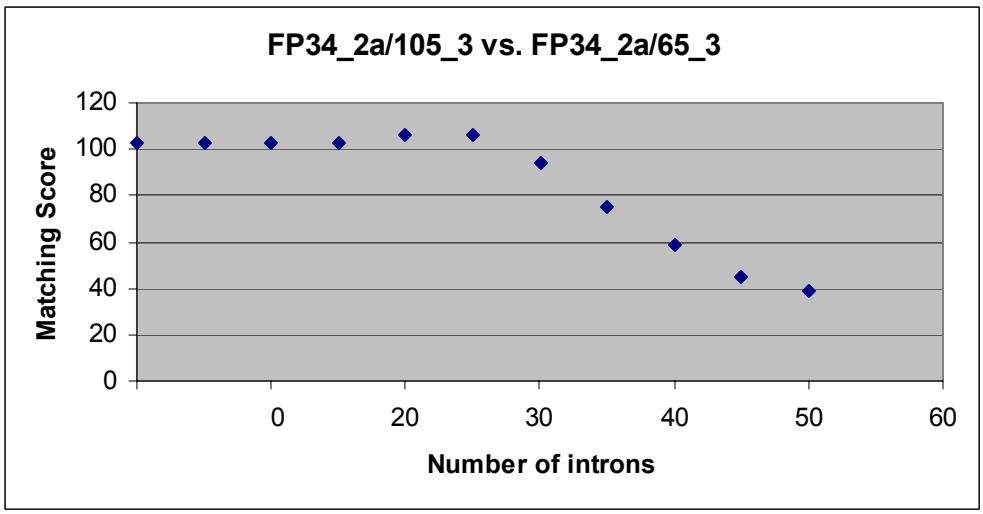

Figure 4.5

From Figure 4.5 we can see that one fingerprint template inserted with one set of introns (From FP1_1) can still match itself inserted with a different set of introns (From FP105_3 and 65_3).

4.1.3 Similar fingerprints with same set of introns

As shown in Figure 4.6, adding the same set of introns into FP34_2a and 34_2b increases the matching scores and so will increase the false matching rate.

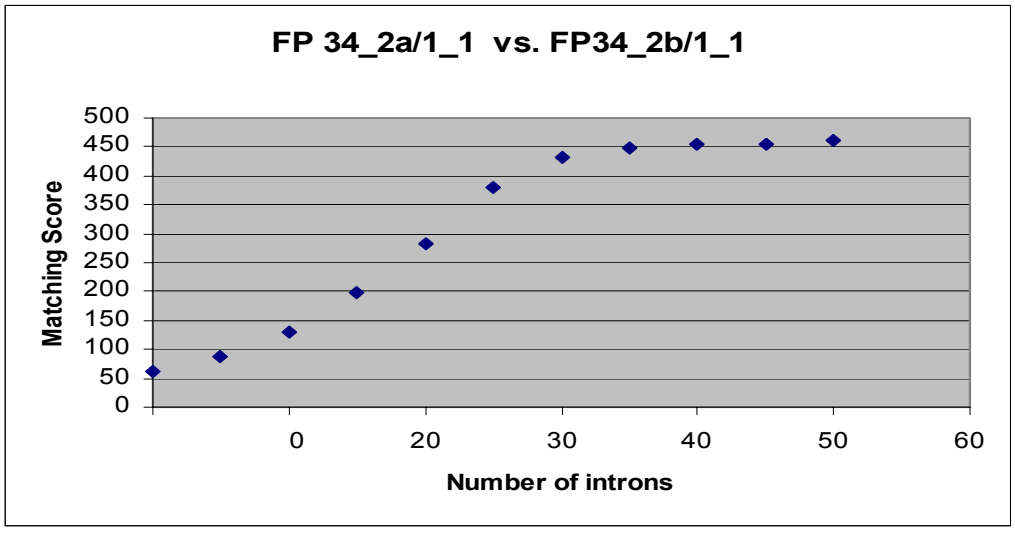

Figure 4.6

4.1.4 Similar fingerprints with different sets of introns

Adding different sets of introns to FP34_2**a** and FP34_2**b** gives the results shown in Figure 4.7 and 4.8.

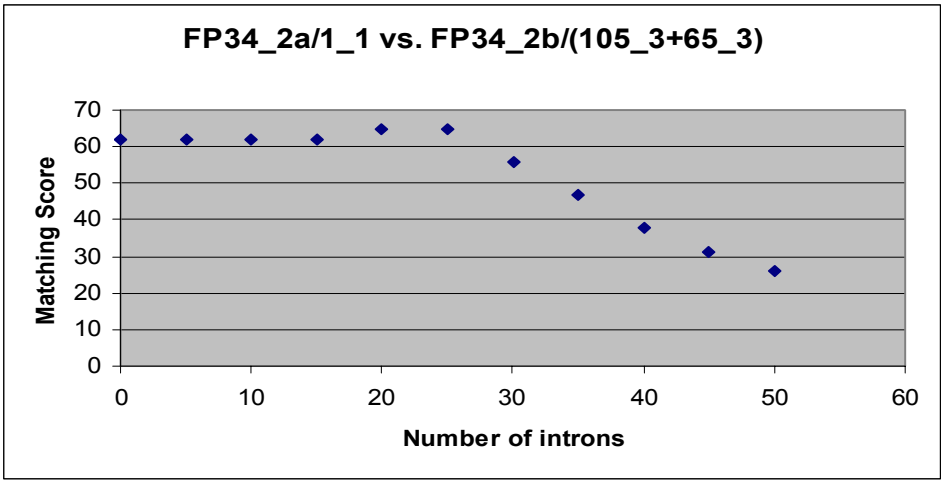

Figure 4.7

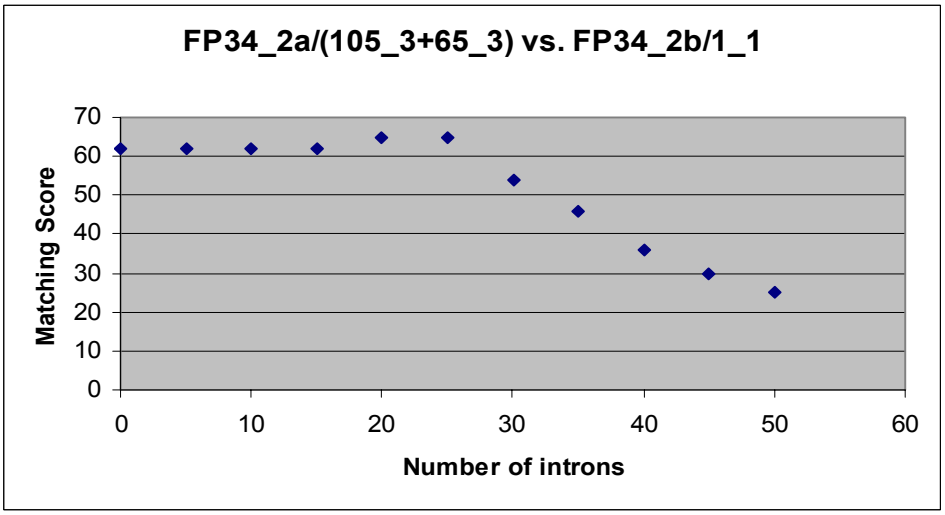

Figure 4.8

For Figure 4.7 and 4.8, the intron sets are taken from FP105_3 first and then from FP65_3. Note that not all the minutiae of FP65_3 are used.

Adding different sets of introns to FP34_2**a** and FP34_2**c** gives the results shown in Figure 4.9 and 4.10.

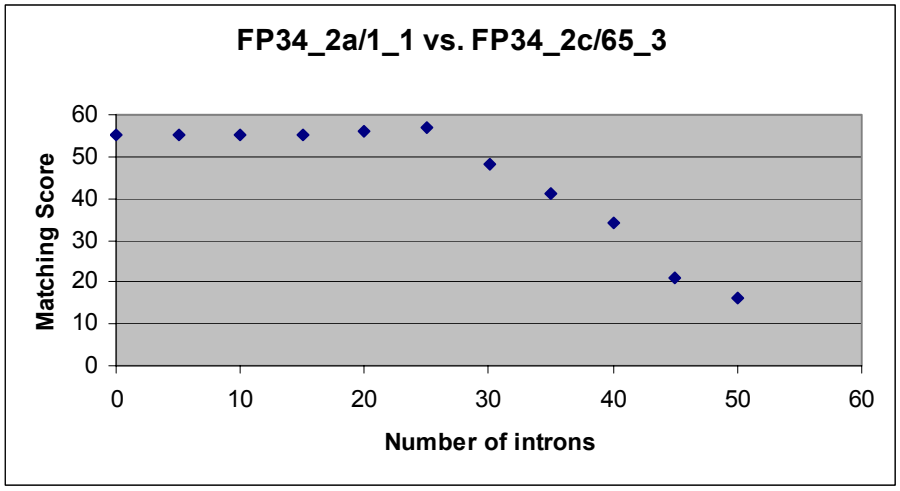

Figure 4.9

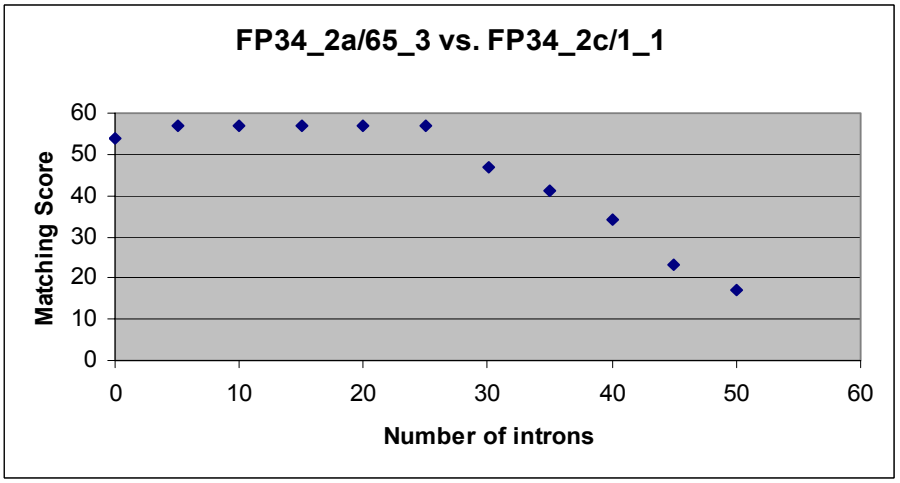

Figure 4.10

The results from Figure 4.7 to 4.10 show that similar fingerprints inserted with different sets of introns can still match, which provides a solution to the non-reproducible problem.

4.1.5 Different fingerprints with same set of introns

For different fingerprints, adding same set of introns will significantly increase the false match rate, as shown in Figure 4.11.

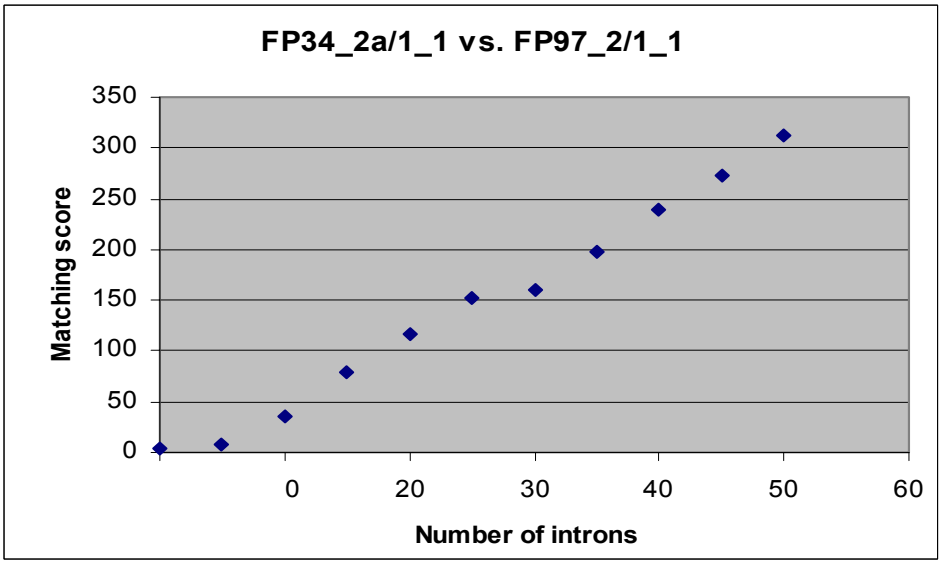

Figure 4.11

From Figure 4.6 and 4.11, we concluded that it should be avoided adding same set of introns.

4.1.6 Different fingerprints with different sets of introns

Two different fingerprints and two different sets of introns are used. The results are given in Figure 4.12 and 4.13.

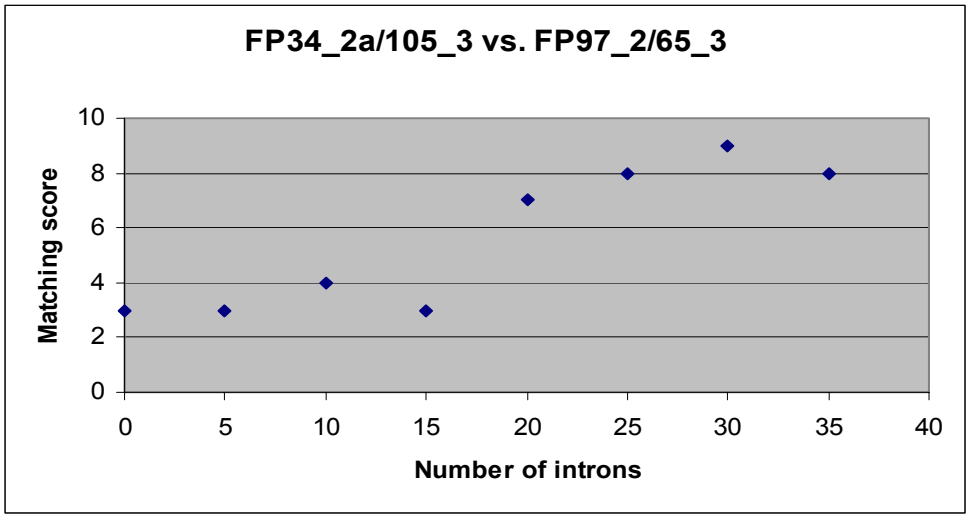

Figure 4.12

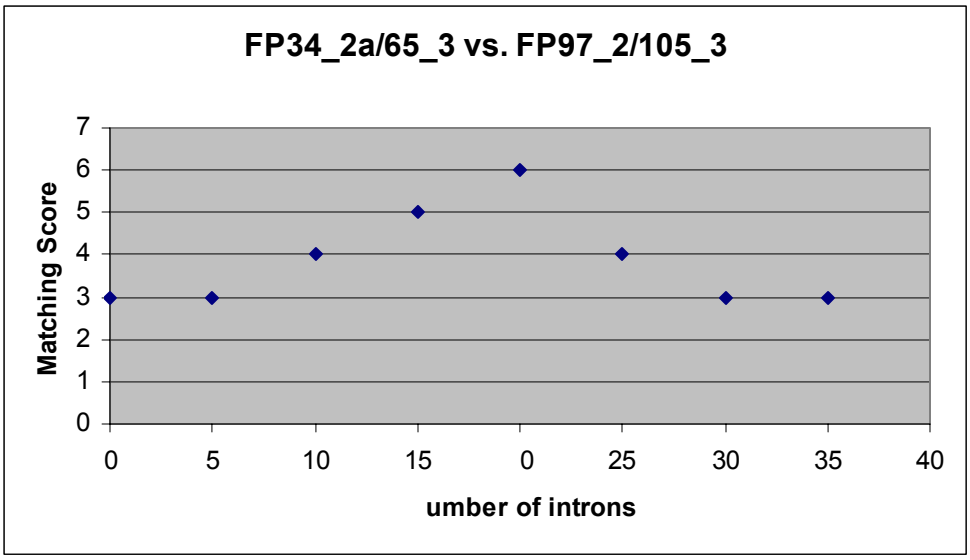

Figure 4.13

From Figure 4.12 and 4.13, we can see that adding different sets of introns to different fingerprints will not significantly change the matching scores.

In sum, the results from Figure 4.1 to 4.13 support the following conclusions:

- AIM can be used as a hashing mechanism for protecting biometrics.
- Different sets of introns should be used for same or similar fingerprints.
- Increasing the similarity of two fingerprint templates may allow a larger Message Expansion Rate.
- Avoid using the same sets of introns.
- Adding different sets of introns to different fingerprints will not significantly change the matching score.

Based on these results, we choose to test adding different sets of introns to different fingerprints as given in section 4.2.

4.2 Identification (1:N)

Depending on where to add introns, probe fingerprint or database, three situations are considered.

4.2.1 Intronize database only

All the intron sets are different. Figure 4.14 shows the results.

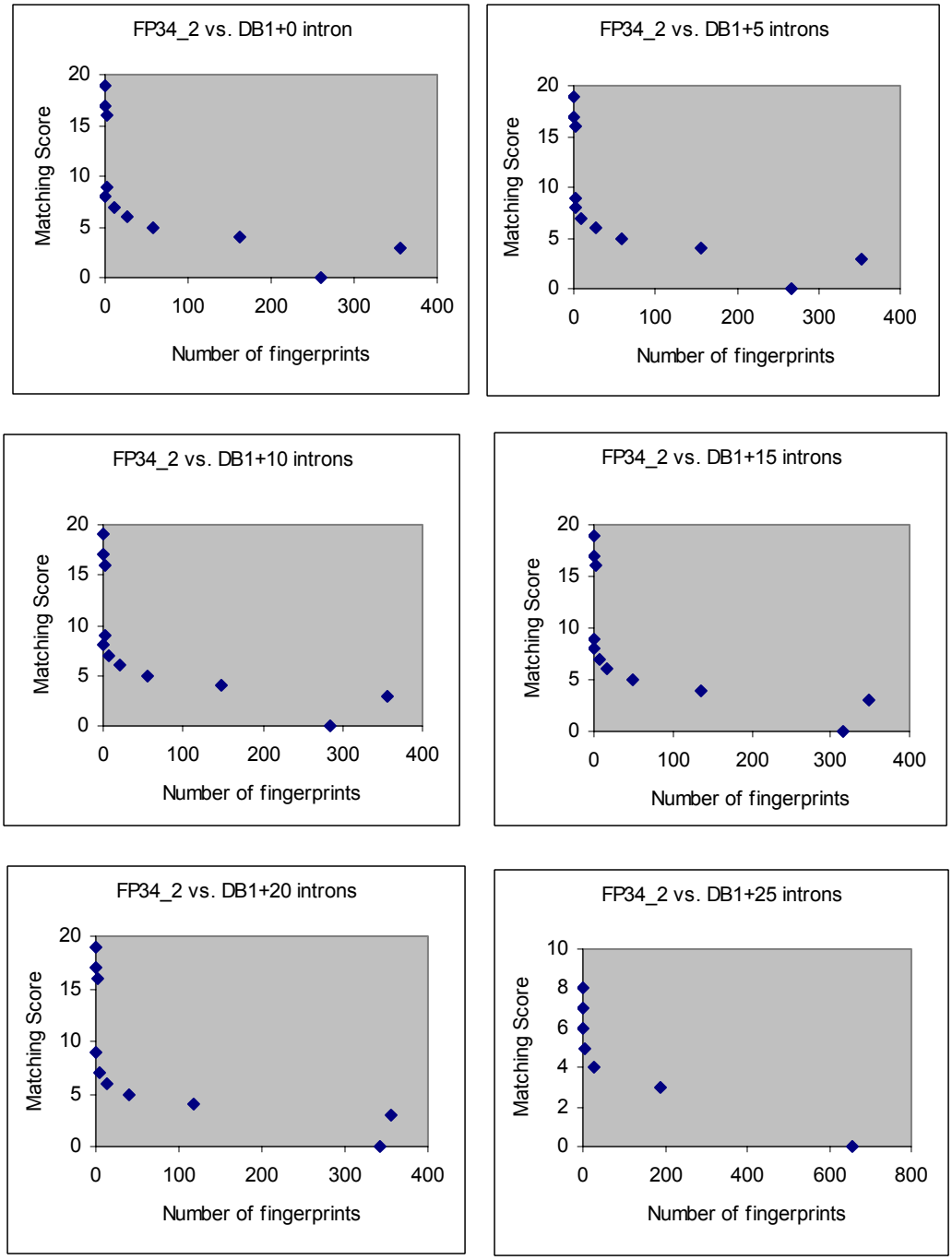

Figure 4.14 Intronize database only

In Figure 4.14, every graph follows a similar distribution with a peak matching score smaller than 5, and none of the matching scores is greater than 20. According to threshold 40 set by NIST [72], the false matching rate is 0.

In database DB1, there are 880 fingerprints. Figure 4.14 only shows 879 of them. The matching score for the probe fingerprint 34_2 against itself is given in Table 4.2.

Table 4.2 Matching scores for probe fingerprint FP34_2

# Introns	0	5	10	15	20	25
Matching Score	103	103	103	103	103	41

Table 4.2 tells us that the false non-matching rate is 0 with up to 25 introns (MER=2). Most importantly, the data in Table 4.2 follows similar pattern as those shown in Figure 4.1 and 4.2, which supports the validity of the testing results.

4.2.2 Intronize probe fingerprint only

The sets of introns for probe fingerprint FP34_2 are obtained from FP1_1. We modified the database DB1 by removing the fingerprints that are related to the probe fingerprint, including 16 fingerprints from FP1_1 to 1_8 and FP34_1 to 34_8. Therefore only 864 fingerprints are left for testing. The results are given in Figure 4.15.

From the matching scores represented by the Y-coordinate, we can see that the peak values are around 5 and the maximum values are less than 20. Therefore, the false matching rate is 0.

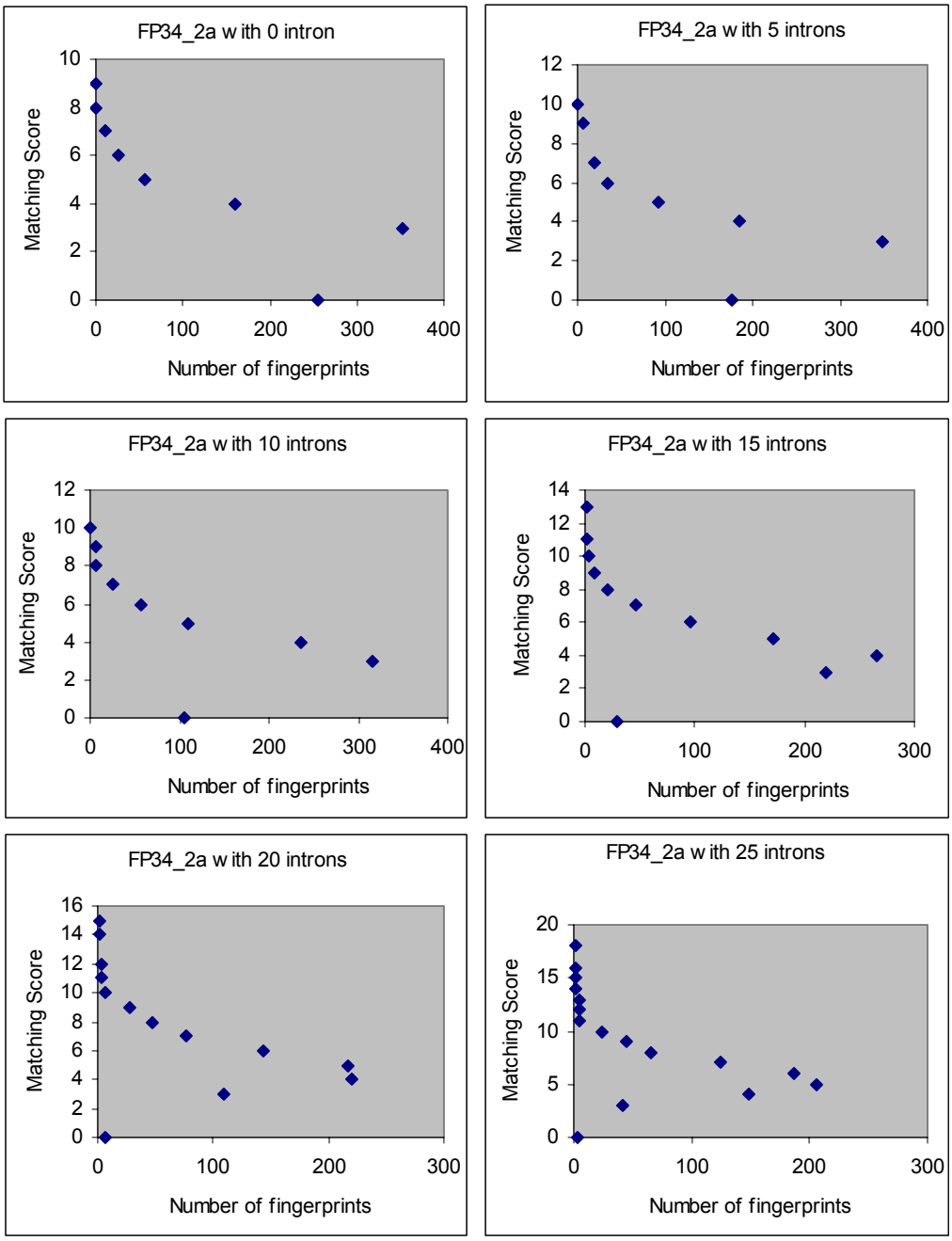

Figure 4.15 Intronize probe fingerprint only

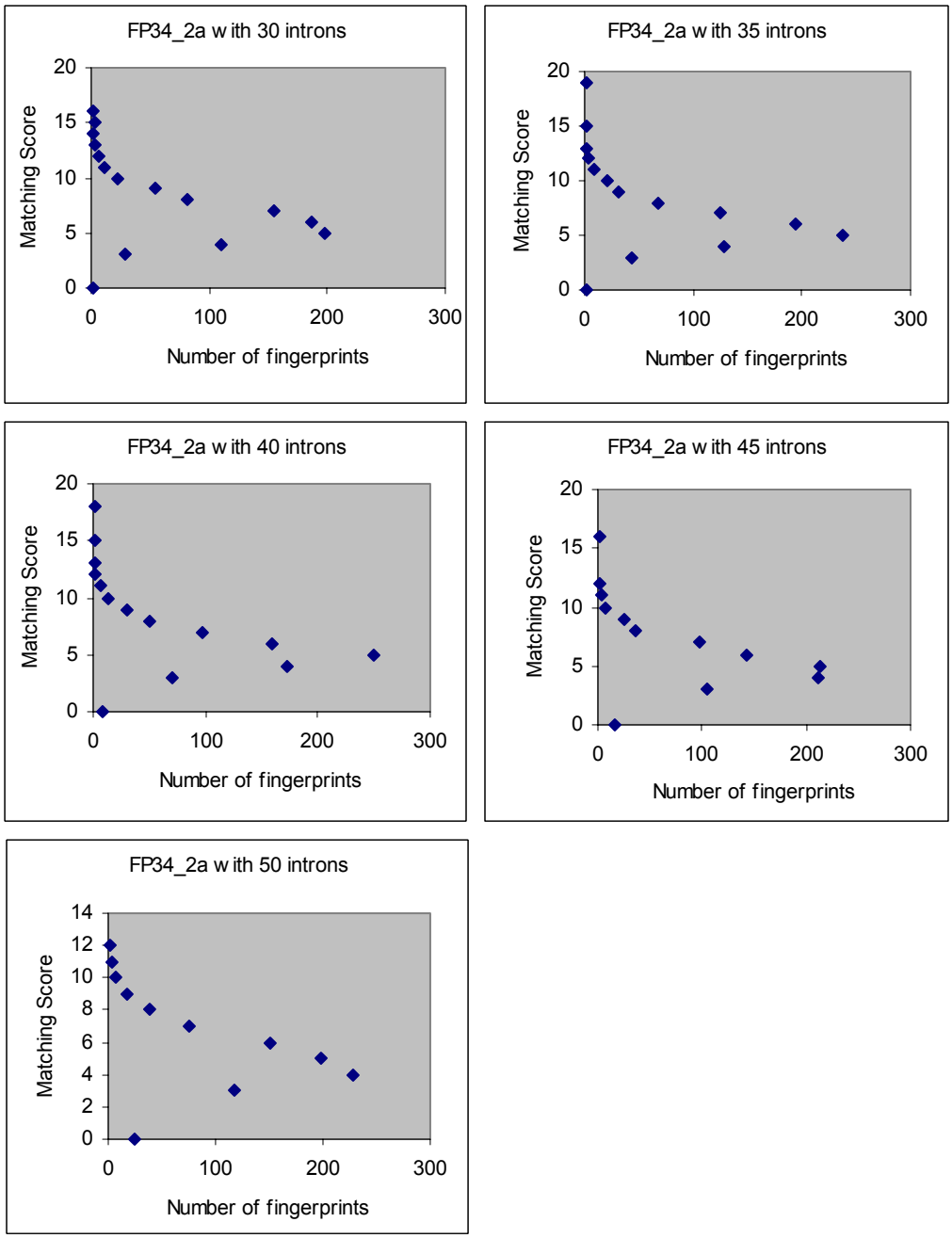

Figure 4.15(continued)

4.2.3 Intronize both probe fingerprint and database

There are 879 fingerprints. All intron sets are different.

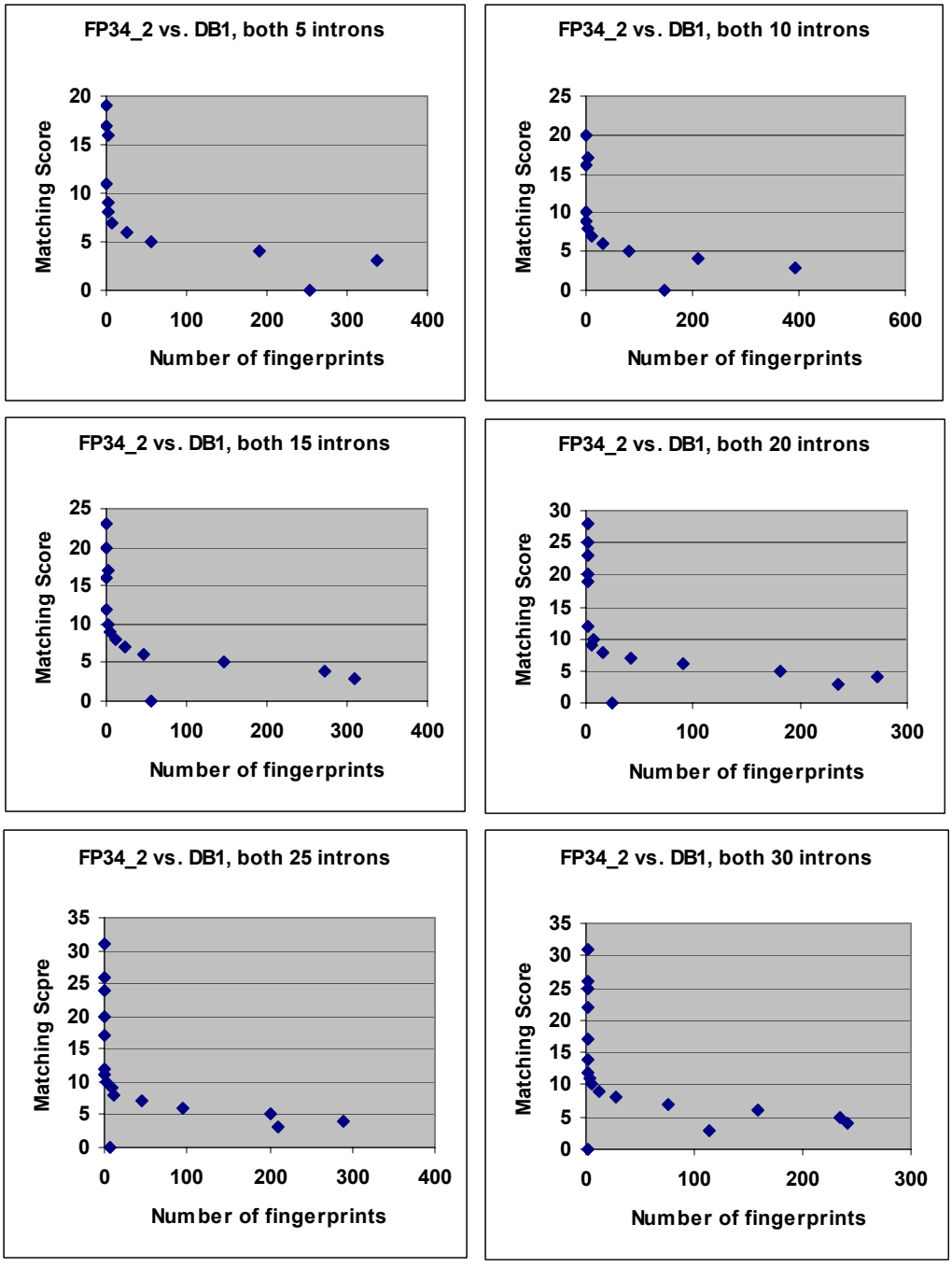

Figure 4.16 Intronize both probe fingerprint and database

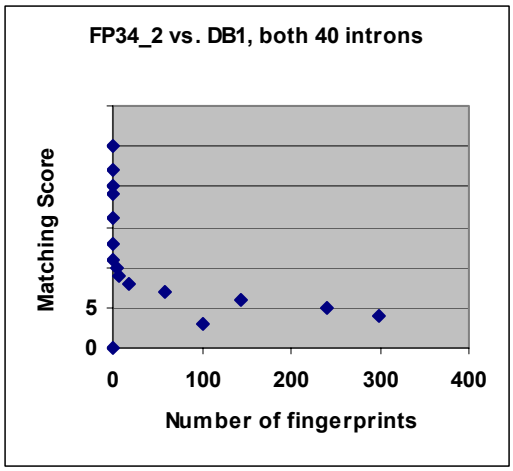

Figure 4.16(continued)

Figure 4.16 shows the testing results as we add up to 40 introns (MER> 2) to both probe fingerprint and database fingerprints.

Table 4.3 gives the matching results of FP34_2 against FP34_1 to 34_8 during the intronization process.

Table 4.3 Matching results for FP34_2 inserted up to 40 introns

	0	5	10	15	20	25	30	40
34_1	4	4	4	0	8	5	5	5
34_2	103	148	147	192	241	320	320	499
34_3	17	17	17	17	28	31	31	25
34_4	9	9	10	10	10	10	10	5
34_5	16	16	16	23	23	24	25	19
34_6	19	19	20	20	20	20	22	22
34_7	16	16	17	17	25	26	26	16
34_8	3	11	5	16	19	17	17	20

From Figure 4.16 and Table 4.3, we can see both the false match rate and the false non-match rate are equal to 0%. These results again show that the AIM can be used as an effective hashing method for protecting fuzzy biometrics.

Chapter 5 Conclusions and Future Work

5.1 Research Contribution

Biometric information needs to be protected. Due to the noisy nature of biometric measurements, widely used hashing algorithms, for example SHA-1and MD-5, do not work. In this thesis, a new method — Artificial Intronization Method is proposed and developed to mainly approach the non-reproducible problem. Our extensive testing results with fingerprint database show that AIM can be used as a hashing mechanism — matching in encrypted formats.

5.2 Future Research

We concluded the thesis by listing the possible future research.

1. More advanced intronization techniques

Nature has created many algorithms that deserve computer scientists to explore. Two terms from Genetics, Alternative Splicing and Restriction Enzymes, could be starting points for advanced intronization algorithm design.

2. Better matching algorithm design

The matching algorithm used in the thesis is based on line segments. Different matching algorithm can be designed, e.g., triangular matching.

Appendix: Biological One-Way Function

Abstract: Two ciphers are involved in the Central Dogma of Biology: intronization cipher and substitution cipher. In Cryptography, substitution has been used widely long before humans recognized DNA; it seems that intronization has not been explicitly proposed as an encryption technique to date, as far as we know.

A.1 Introduction

One-way function plays an essential role for information security. Inspired by the existing Mathematical One-Way Function [A1] and Physical One-Way Function [A2-18], we introduce a new term: Biological One-Way function.

A.2 Biological One-Way Function

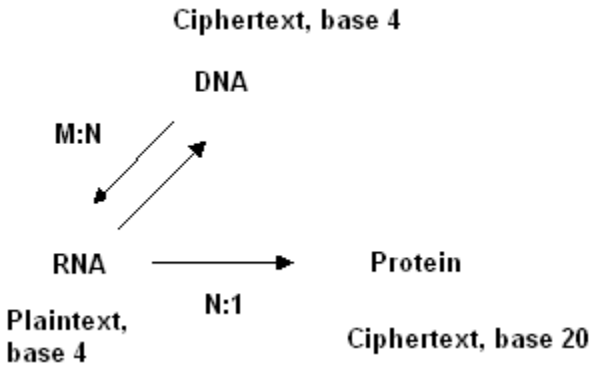

Figure A.1 The Central Dogma of Biology

The Central Dogma of Biology, transcription of DNA to RNA and translation from RNA to protein can be described with the five steps [A19]:

1) DNA replication

2) DNA is transcribed into precursor-mRNA(pre-mRNA) → Transcription

3) pre-mRNA is spliced to form messenger RNA(mRNA)

4) mRNA moves from nucleus to cytoplasm.

5) mRNA is translated into protein → Translation

The one-way-ness of the protein production process is represented with two ciphers.

The first cipher is the splicing of introns from the pre-mRNA transcript – we call it Intronization Cipher.

In the eukaryotic cell, only less than 10% of the entire DNA sequence is directly used for protein coding, that is to say, a large amount of non-coding regions exists in DNA. Modern biologists believe that these non-coding regions have yet-to-be-known functions. Where did the redundancy come from? That evolution selects and modifies genetic sequences along the way may not account for the origination of all the introns. Two theories, named Intron Early and Intron Late, exist.

From cryptographic perspective, we believe the protection due to the existence of the non-coding regions of DNA helps organisms to survive.

Just given a genomic DNA segment it is difficult to predict where the splicing sites are. However, some knowledge has been gathered through biochemical experiments, for example:

•The length of mRNA is a multiple of 3

•5' splice sites (exon to intron) are usually GT

•3' splice sites (intron to exon) are usually AG

With this knowledge, genes can be recognized from DNA sequences using similarity search and statistical analysis. However, difficulties still exist to accurately predict

genes. For example, Alternative Splicing - same gene splices differently and produces different proteins, does exist [A21-25]. The difficulty of predicting the splicing is the first reason we call it Biological One-Way Function.

The second cipher comes from the mapping of mRNA sequence to protein sequence. As [A20] pointed out, "*the genetic code is a substitution cipher*".

In the genetic code, there are 61 codons (triplets of 4 letters A, U, G, C, 4x4x4) coding for 20 amino acids and 3 codons signaling the stop of translation. Except the 1-to-1 correspondences between amino acid Tryptophan and codon UGG, there are 2, 3, 4, 5 or 6 codons coding for each individual amino acid. On average the number of possible RNA sequences for a given protein sequence of length n will be 3^n. Due to the redundancy of codons, it is difficult to reversely transcribe a protein sequence into RNA sequence. That is the second reason we called it Biological One-Way Function.

Is it safe to directly use the genetic code to encrypt English messages? The answer is NO because the 3-lettered codons generally start with two same letters in the same order. Therefore, on average reverse translation of a protein sequence could correctly reproduce about two thirds of its RNA sequence. With the redundancy of English language, ciphertext can be decrypted easily. Therefore, it is necessary to randomize the mapping of the natural genetic code or design new mapping to utilize the one-way-ness.

Biological key

Encoding and decoding of biological information requires a key or a set of keys, which include RNA molecules, proteins, and enzymes, among other things. External environments also play some roles of a key. One example is that the eggs of

crocodiles might hatch into male or female, depending on the environmental temperatures.

Attack

The weakest link expressed in the Central Dogma of Biology is the intermediate code, mRNA, which can be viewed as plaintext. One example of a biological attack starting with RNA is the HIV virus [A26].

Bibliography

Chapter 1

[1] S.C. Draper, A. Khisti, E. Martinian, A.Vetro, and J.S. Yedidia. "Using Distributed Source Coding to Secure Fingerprint Biometrics", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 2, pp. II-129--II-132, April 2007

[2] S.C. Draper, A.Khisti, E.Martinian, A.Vetro, and J.S.Yedidia. "Secure Storage of Fingerprint Biometrics Using Slepian-Wolf Codes", *Information Theory and Applications Workshop (ITA)*, January 2007.

[3] Umut Uludag. "Secure Biometric Systems", PhD Dissertation, Michgan State University (2006).

[4] Feng Hao, Ross Anderson, and John Daugman. "Combining cryptography with biometrics effectively", Technical reports, University of Cambridge, Computer Laboratory (Jul 2005).

[5] US Patent 20020112177, "Anonymous biometric authentication ". Publication date: 08/15/2002

[6] Y. Chen and A. K. Jain. "Dots And Incipients: Extended Features for Partial Fingerprint Matching", *Proc. Biometric Symposium, BCC*, Baltimore (September, 2007).

[7] Nalini K. Ratha, Sharat Chikkerur, Jonathan H. Connell, Ruud M. Bolle. "Generating Cancelable Fingerprint Templates*", IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 29, No. 4, pp. 561-572 (April 2007).

[8] N. K. Ratha, J. H. Connell, and R. M. Bolle. “Enhancing security and privacy in biometrics-based authentication systems.” *IBM Systems Journal*, 40(2): 614--634, 2001.

[9] Nalini Ratha, Jonathan Connell, Ruud M. Bolle, Sharat Chikkerur. "Cancelable Biometrics: A Case Study in Fingerprints," *ICPR*, pp. 370-373, 18th International Conference on Pattern Recognition (ICPR'06).

[10] Marios Savvides, B. V. K. Vijaya Kumar, Pradeep K. Khosla. “Cancelable Biometric Filters for Face Recognition”. *ICPR* (3) 2004: 922-925

[11] T. Boult. “Robust Distance Measures for Face-Recognition Supporting Revocable Biometric Tokens”**,** *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition* (FGR06).

[12] P. Tuyls and Jasper Goseling. “Capacity and Examples of Template Protecting Biometric Authentication Systems”, *BioAW2004*, Praag.

[13] P. Tuyls, E. Verbitskiy, T. Ignatenko, D.W.E. Schobben, A.H.M. Akkermans. “Privacy Protected Biometric Templates: Acoustic Ear Identification”, *SPIE Defense and Security Symposium*, April 2004, Orlando.

[14] P. Linnartz, P. Tuyls. “New Shielding Functions to Enhance Privacy and Prevent Misuse of Biometric Templates”, *AVBPA* 2003, *LNCS*.

[15] E. Verbitskiy, P. Tuyls, D. Denteneer, J.P. Linnartz. “Reliable Biometric Authentication with Privacy Protection*”*, *Benelux Symposium on Information Theory*, May 2003.

[16] P. Tuyls, E. Verbitskiy, J. Goseling, D. Denteneer. "Privacy Protecting Biometric Authentication Systems: An Overview", Philip Research (2004). Available at: http://www.eurasip.org/content/Eusipco/2004/defevent/papers/cr1921.pdf

[17] J. Goseling, P. Tuyls. "Information Theoretic Approach to Privacy Protection of Biometric Templates", *ISIT 2004.*

[18] Russell Ang, Reihaneh Safavi-Naini, Luke McAven. "Cancelable Key-Based Fingerprint Templates", *Australasian Conference on Information Security and Privacy* (2005).

[19] A. Teoh and D. Ngo. "Cancelable Biometrics Featuring With Tokenised Random Number". *Pattern Recognition Letter*, Elsevier Science (2004).

[20] W. Yip, A. Goh, D. Ngo, A. Teoh. "Generation of Replaceable Cryptographic Keys from Dynamic Handwritten Signatures*"*. *LNCS*(2006).

[21] W. Yip, A. Goh, D. Ngo, A. Teoh. "Refreshable Cryptographic Key Generation from Dynamic Handwritten Signature*"*. *Special issue on Pattern Recognition in Biometrics and Bioinformatics, Multimedia CyberScape Journal* (2005).

[22] Ying-Han Pang, Andrew Teoh, David Ngo. "Cancelable Palmprint Authentication System*"*. *International Journal of Signal Processin*, Vol. 1, No. 2, pp.98-104 (2004).

[23] Neo Han and Andrew Teoh. "A Novel Cancelable Face Authentication Biometrics". *Special issue on Pattern Recognition in Biometrics and Bioinformatics, Multimedia CyberScape Journal* (2005).

[24] Ying-Han Pang, Andrew Teoh, and David Ngo. "Binarized Revocable Biometrics in Face Recognition". *LNCS* (2005).

[25] Andrew Teoh, Tee Connie. “Remarks on BioHashing based Cancelable Biometrics in Verification System.” *NeuroComputing* (2006), Elservier Science.

[26] Andrew Teoh, David Ngo and Alwyn Goh. “Biohashing: Two Factor Authentication Featuring Fingerprint Data And Tokenised Random Number”. *Pattern Recognition*, Vol 37, Issue 11, pp 2245-2255, Elservier Science.

[27] Andrew Teoh, David Ngo and Alwyn Goh. “Personalised Cryptographic Key Generation Based On Facehashing”. *Computer and Security* (2003), Elsevier Science, UK

[28] Tee Connie, Andrew Teoh, Michael Goh, and David Ngo. “PalmHashing: A Novel Approach for Dual Factor Authentication”, *Pattern Analysis and Applications* (2003*)*, Springer-Verlag, UK.

[29] Tee Connie, Andrew Teoh, Michael Goh and David Ngo. “PalmHashing: A Novel Approach for Cancelable Biometrics”, *Information Processing Letter* 2003, Elsevier Science, UK

[30] A. Lumini and L. Nanni. “An improved BioHashing for human authenticaiton”, *Pattern Recognition* (2006)

[31] L. Nanni and A. Lumini. "Empirical Tests on BioHashing", *NeuroComputing*, vol.69, no.16, pp.2390-2395, October 2005.

[32] D. Maio and L. Nanni. "MultiHashing, human authentication featuring biometrics data and tokenised random number: a case study FVC2004", *NeuroComputing*, vol.69, pp.242-249, December 2005

[33] G. Davida, Y. Frankel, B. Matt. “On enabling secure applications through off-line biometric identification”, *Proceedings Symposium on Privacy and Security*, 1998, pp 148-157.

[34] M. van der Veen, T. Kevenaar, G. Schrijen, A. Akkermans, and F. Zuo. “Face biometrics with renewable templates”, *Proceedings of SPIE* -- Volume 6072 *Security, Steganography, and Watermarking of Multimedia Contents VIII* (Feb. 15, 2006)

[35] T. Kevenaar, G. Schrijen, M. van der Veen, A. Akkermans, F. Zuo. “Face recognition with renewable and privacy preserving binary templates”, *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, 2005: 21- 25.

[36] M. Indovina, U. Uludag, R. Snelick, A. Mink and A. Jain. "Multimodal Biometric Authentication Methods: A COTS Approach", *Proc. MMUA 2003, Workshop on Multimodal User Authentication*, pp. 99-106, Santa Barbara, CA, December 11-12, 2003.

[37] Y. Wang, T. Tan and A. K. Jain. "Combining Face and Iris Biometrics for Identity Verification", *Proc. of 4th Int'l Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pp. 805-813, Guildford, UK, June 9-11, 2003.

[38] A. Kumar, D. C. M. Wong, H. C. Shen and A. K. Jain. "Personal Verification Using Palmprint and Hand Geometry Biometric", *Proc. of 4th Int'l Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pp. 668-678, Guildford, UK, June 9-11, 2003.

[39] A. Ross and A. K. Jain, "Information Fusion in Biometrics", *Pattern Recognition Letters*, Vol. 24, Issue 13, pp. 2115-2125 (September, 2003).

[40] A. Ross, A. K. Jain, and Jian Zhong Qian. "Information Fusion in Biometrics", *Proc. 3rd International Conference on Audio- and Video-Based Person Authentication (AVBPA),*, pp. 354-359, Sweden, June 6-8, 2001.

[41] A. K. Jain, S. Prabhakar and S. Chen. " Combining Multiple Matchers for a High Security Fingerprint Verification System", *Pattern Recognition Letters*, Vol 20, No. 11-13, pp. 1371-1379, 1999.

[42] L. Hong, A. Jain and S. Pankanti. "Can Multibiometrics Improve performance?", *Proceedings AutoID'99*, Summit, NJ, Oct 1999, PP. 59-64.

[43] A.K. Jain, L.Hong, Y. Kulkarni. "A Multimodal Biometric System using Fingerprints, Face and Speech", *2nd Int'l Conference on Audio- and Video-based Biometric Person Authentication*, Washington D.C., pp. 182-187, March 22-24, 1999.

[44] L. Hong and A.K. Jain. " Integrating Faces and Fingerprints For Personal Identification", *IEEE Transactions PAMI* , Vol.20, No.12, pp 1295-1307, 1998.

[45] A.K. Jain, L. Hong, and Y. Kulkarni. "F2ID: A Personal identification System Using Faces and Fingerprints", *Proc. 14th Int'l. Conf. Pattern Recognition*, Brisbane, pp. 1373 - 1375, August 1998.

[46] L. Hong and A.K. Jain. "Integrating Faces and Fingerprints for Personal Identification", *Proc. 3rd ACCV*, Hong Kong, Jan. 1998.

[47] L. Hong and A. Jain. "Automatic Personal Identification by Integrating Faces and Fingerprints", *Proc. IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 15-18, Stony Brook, NY, November, 1997.

[48] US Patent 7120607, "Business system and method using a distorted biometrics". Pubication Date: 10/10/2006.

[49] US Patent 6735695, "Methods and apparatus for restricting access of a user using random partial biometrics". Publication date: 05/11/2004.

[50] Meredith Wadman. "Biometrics group counters privacy fears". *Nature*, Vol 398, Issue 6727, p.451

[51] B. Schneier. "Inside risks: The use and abuse of biometrics". *Communications of the ACM*, Vol. 42, pp.136.

[52] Christopher Coelle. "The Use and Abuse of Iris Recognition Technology". Available at: http://www.coelle.org/papers/TheUseandAbuse.pdf

[53] Jan Grijpink. "Two barriers to realizing the benefits of biometrics: a chain perspective on biometrics and identity fraud as biometrics' real challenge", *Proceedings of SPIE* -- Volume 5310. Optical Security and Counterfeit Deterrence Techniques V, Rudolf L. van Renesse, Editor, June 2004, V.5310, pp. 90-102

[54] M. Bronstein and A. Bronstrein. "Biometrics was no match for hair-raising tricks", *Nature*, Volume 420, Issue 6917, pp. 739 (2002).

[55] Ileana Buhan, Pieter Hartel. "The State of the Art in Abuse of Biometrics". Available at: http://www.coelle.org/papers/TheUseandAbuse.pdf

[56] William Abernathy and Lee Tien. “Biometrics: Who's Watching You”, *Electronic Frontier Foundation.*

Available at: http://www.eff.org/Privacy/Surveillance/biometrics/

[57] Qinghai Gao, Xiangdong Li, Lin Leung and Michael Anshel. “Survey of Biometric Application on Cryptography”, *International Workshop on Advanced Image Technology* (IWAIT), Bangkok Conference, Thailand, Jan.8, 2007.

Chapter 2

[58] Claude E. Shannon. "Communication Theory of Secrecy Systems", *Bell System Technical Journal*, Vol.28-4, page 656--715, 1949.

[59] Available at: http://awards.acm.org/

[60] N. J. A. Sloane. Available at :http://www.research.att.com/~njas/sequences/

[61] N. J. A. Sloane. “The on-line encyclopedia of integer sequences”. *Notices of the AMS*, Vol. 50, No.8, Sep., 2003.

[62] Available at: http://ja0hxv.calico.jp/pai/epivalue.html

[63] Annie S. Wu, Robert K. Lindsay. “A Survey of Intron Research in Genetics”, *Proceedings of the 4th International Conference on Parallel Problem Solving from Nature, Lecture Notes In Computer Science*; Vol. 1141, p101-110 (1996).

Chapter 3

[64] International Biometric Group. “Generating Images from Templates”, IBG White Paper, 2002. Available at:

http://www.biometricgroup.com/reports/public/reports/templates_images.html

[65] C.J. Hill. “Risk of Masquerade Arising from the Storage of Biometrics”, Thesis for the Bachelor of Science, Dept. of CS, Australian National University, 2002

[66] A. Adler. “Sample images can be independently restored from face recognition templates*”, Canadian Conference on Electrical and Computer Engineering*, May 2003.

[67] A. Ross, J. Shah and A. K. Jain. "From Template to Image: Reconstructing Fingerprints From Minutiae Points," *IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Biometrics,* Vol. 29, No. 4, pp. 544-560, April 2007.

[68] A. Ross, J. Shah, and A. K. Jain. "Towards Reconstructing Fingerprints from Minutiae Points", *Proc. of SPIE Conference on Biometric Technology for Human Identification II*, (Orlando, USA), pp. 68-80, March 2005.

[69] US Patent 20040193893, “Application-specific biometric templates”. Publication date: 09/30/2004

[70] U. Uludag, S. Pankanti and A. K. Jain. "Fuzzy Vault for Fingerprints", *Proc. of Audio- and Video-based Biometric Person Authentication* (AVBPA) 2005, pp. 310-319, Rye Brook, NY, July 2005.

Chapter 4

[71] Available online: http://bias.csr.unibo.it/fvc2004/.

[72] NIST Biometric Image Software (NBIS). Available online: http://fingerprint.nist.gov/

Appendix

[A1] Susan Landau. “Find Me a Hash”, *ACM* 2006.

[A2] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld. “Physical One-Way Functions”. *Science*, 297:2026-2030, 2002.

[A3] R. Pappu. “Physical One-Way Functions”. PhD thesis, MIT, 2001.

[A4] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas. “Silicon Physical Random Functions”. *Proceedings of the Computer and Communication Security Conference*, November 2002.

[A5] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas. “Controlled physical random functions”. *Proceedings of 18th Annual Computer Security Applications Conference*, December 2002.

[A6] B. Gassend. Physical Random Functions. MS thesis, MIT, 2003.

[A7] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas. “Delay-based circuit authentication and applications”. *Proceedings of the 2003 ACM Symposium on Applied Computing*, March 2003.

[A8] Daihyun et al. “Extracting secret keys from integrated circuits”, *IEEE Transactions on VLSI Systems (2005).*

[A9] M. Dijk, D. Lim and S. Devadas. “Reliable Secret Sharing With Physical Random Function”, *Tech. Rep.,* MIT Computation Structures Group MEMO-475, May 2004.

[A10] J. W. Lee et al. “ A Technique to Build a Secret Key in Integrated Circuits for Identification and Authentication Applications”. *Symposium on VLSI Circuits, 2004.*

[A11] P. Tuyls et al. “Security analysis of Physical Uncloneable Functions*”, 25th Benelux Symposium*, 2004.

[A12] P. Tuyls et al. “An information theoretic model for physical uncloneable functions**”**, *Proceedings of the International Symposium on Information Theory* (2004).

[A13] P. Tuys et al. “Information-Theoretic Security Analysis of Physical Uncloneable Functions,” *Proc. 9th Conf. on Financial Cryptography and Data Security* (2005).

[A14] B. Škorić, P. Tuyls, W. Ophey. “Robust key extraction from Physical Uncloneable Functions,” *Proc. Applied Cryptography and Network Security (ACNS) 2005*.

[A15] B. Škorić, S. Maubach, T. Kevenaar, P. Tuyls. “Information-theoretic analysis of capacitive Physical Uncloneable Functions,” *J. Appl. Phys*. 100, 024902 (2006).

[A16] P. Tuyls et al. “Read-proof hardware from protective coatings,” *CHES* 2006.

[A17] P. Tuyls and B. Škorić. “Physical Uncloneable Functions for Enhanced Security of Tokens and Tags”, *ISSE* 2006.

[A18] N. Bird, C. Conrado, J. Guajardo, S. Maubach, G.J. Schrijen, B. Škorić, P. Thueringer, A.M.H. Tombeur, P. Tuyls. “ALGSICS - Combining Physics and Cryptography to Enhance Security and Privacy in RFID Systems”, *WISSEC* 2006.

[A19] Available at: http://www.accessexcellence.org/

[A20] Mark White. “Geometric Structure of Codon Relationships”, Published online (2004). Available at: http://www.codefun.com/.

[A21] A. Matlin, F. Clark, C. Smith. “Understanding alternative splicing: towards a cellular code", *Nat. Rev. Mol. Cell Biol*. 2005, 6:386-398.

[A22] C. Lee and Q. Wang. "Bioinformatics analysis of alternative splicing", *Brief Bioinform. 2005, 6: 23-33.*

[A23] S. Stamm et al. "Function of alternative splicing", *Gene* 2005, 344: 1-20.

[A24] R. Sorek, R. Shamir, G. Ast. "How prevalent is functional alternative splicing in the human genome?" *Trends Genet*. 2004, 20: 68-71.

[A25] L. Lareau, R. Green, R.Bhatnagar, S.Brenner. "The evolving roles of alternative splicing", *Curr. Opin. Struct. Biol*. 2004, 14: 273-282.

[A26] http://www.ishipress.com/aids.htm