Design and Implementation of E-Document Encryption System using Hash Algorithm

Jung-Oh Park,
Dept of Computer Science, Soongsil University
Jop07@ssu.ac.kr

Sang-Geun Kim
Division of Computer Engineering, Sungkyul University
sgkim@sungkyul.edu

Abstract

It is increasing to use ID card and passport for identification as development of information communication technology on 21century. But image of human face that uses for e-passport and admission tag is inserted on RFID chip without any encryption process, or is encrypted using general encryption process. These demands are ever increasing in such business circles, as e-trade businesses where various e-documents go through complex routes and the ownerships of the documents are transferred through a number of parties, and logistics, legal and pharmaceutics circles in which the subjects who record, maintain and read e-documents continue changing. The existing encryption algorithms to ensure the originality of e-documents generally encode the whole area of documents with a symmetric key, so if a user discloses the corresponding symmetric key, it becomes impossible to secure the safety of related documents. In addition, the algorithms have a disadvantage that a user should show the entire information including unnecessary parts (unwanted parts) to receivers in case the entire area of a document is encrypted. This study, thus, proposes a scrambling encryption technique, by which a user, while issuing a registered e-document to a third party, can send partial data of the document, not the entire information, in order to prevent unnecessary leaking of information, improve the readability of documents and use a hash function instead of a complicated encryption process in changing values of image pixels to relocate the image. The proposed system is designed, implemented and tested in performance. In the test, the safety and the operation speed of the system are measured, and with test results, the system is verified in validity and superiority..

Keywords: *Electronic ID, Scrambling, Originality, Integrity, Readability*

1. Introduction

Use frequency of certificate that can prove one's ID such as ID card or passport is rapidly increasing as information oriented era is becoming actualized due to development of the 21st century IT technology, and due to this, protection of privacy or personal information is becoming an important issue [4,5,7]. All sorts of services are appearing on the internet for e-commerce but securing of security for orders purchase process about many trades achieved through this and measures on all kinds of information threat and essential certification and security that are necessary for performing payment are not complete yet. Moreover, company with reinforced security or using computer at that place distincs if the user is certified through using ID card and whether a person can access or use is decided through this process. .

One's face data that is used in electronic entering system or e-passport is inserted without any encrypting process or through an encrypting process using regular encrypt key in the

existing studies. In other words, it is safe to use it in general purpose system because encrypt key can't be known but it has to be expanded to a system that decodes password by receiving encrypt key through online in order to solve this, and devices using RFID can't be used due to its lack of calculation ability. Moreover, a system using an identification card used in an offline environment is used by inserting it into a smart chip that can't be modified or it can be falsified at a place where modification is possible so user data can be exposed intactly. The case of using lightweight symmetry algorithm is also unfeasible process in RFID tag.

Therefore the suggesting system proposes a method that encodes and decodes image data used in entering and exiting certificate using the bit scrambling method and a method to certify an user due to comparing the original picture with the scrambling encoded data, The suggesting system uses a simple function like XOR instead of complicated encrypting algorithm like symmetry key or open key and it can be used in RFID system, which has relatively weak performance because hash algorithm is used to produce a key. The hash algorithm is used to extend length of key as in stream encrypting through hash chain method, and the place that creates key using hash function is calculated in the device with more advanced calculation ability like RFID transceiver than RFID tag and then the result is used in communication with tag.

In this study, the suggested system is designed and realized, encrypting and decoding experiments using scrambling method are conducted many times in order to assess performance, and safety of the suggested algorithm is proved..

2. Related Studies

2.1 Technologies of Digital Data Protection

Digital data have an advantage that they make it possible to perform flexible operation and various functions between media, but also have such disadvantages as easy reproduction and manipulation of data and difficulty in drawing a line between replicated and original data. These weaknesses have been considered significant because they can damage intellectual property rights of copyright holders. Technologies to protect digital data are, thus, required to solve those problems.

2.1.1 Steganography

Steganography is a technology to embed a message in meaningless contents in order to hide the existence of the message. Even though this method is similar to watermarking, it is different from watermarking in many aspects.

Steganography is to hide information, in which capacity determines the efficiency of transmission, invisibility is put above robustness, and a public key algorithm is applied.

To hide information, steganography uses such methods as invisible ink, small-dot size pictures, sequence of characters and digital signature. The purpose of steganography is that third parties do not recognize even the presence of a message hidden in image data.

Steganography can be classified by features of information, as shown in Table 1.

Table 1. Classification of Steganographic Systems by Features of Information

	Shared Information	Safety	Attacker Model	Weakness
Pure Steganography	No	Embedment, extraction	Passive	No actual case on weakness
Private-Key Steganography	Private key	Key or cover	Passive	Sharing of the key
Public-Key Steganography	Public key	Key	Active	Difficulty in deciding if a hidden message is existent

2.1.2 Watermarking

Watermarking schemes are to transform a work in a way that it is impossible to distinguish the work with the eye, so as to keep its message [9]. In other words, additional information such as license or ownership is inserted to original contents in order to protect intellectual property rights and prevent illegal circulation and manipulation of information.

A watermarking system generally consists of an embedder and a detector. The embedder takes a message or mark as a watermark to be input. The value computed from the embedder is taken as an input of the watermark detector. Most of the detectors decide whether a watermark is present and if there is an output value embedded with the message.

2.1.3 Scrambling

Scrambling is a method that modifies or encodes original image data with a specific key and transmits the data, only to receivers with an appropriately set key, and allows them to restore the data. Only those with a key authorized to restore original image data can turn the image distorted through scrambling to an original one, and receivers not authorized, even in case decrypting the received image data, receive an image distorted through scrambling. Scrambling systems are, thus, able to protect the right of legitimate receivers. Scrambling methods can be divided largely into scrambling in the spatial domain, scrambling in the frequency domain, scrambling with motion vectors and scrambling with an encryption algorithm [3].

1) Scrambling in the Spatial Domain

It is a common method of scrambling that distorts directly image data on the spatial domain, i.e. a screen. The scrambling techniques in the spatial domain include line reversal, line inversion, and cut and rotate scrambling methods.

- Line Reversal Scrambling: It is the simplest among scrambling methods. Scan lines are turned over: the right (R) side of a scan line is turned to the left (L) and the left side of the line to the right. Reversing the order of transmission of digital data also is a method of line reversal scrambling. This is simple to implement, but is very weak in the security level.

- Line Inversion Scrambling: The values of luminosity for each of the scan lines are obtained, through which light areas become dark and vice-versa. By reversing the values of luminosity, it is possible to distort the whole image or some of scan lines. Even though the method is simple to implement, it is less capable in quality, image distortion and security.

2) Scrambling in the Frequency Domain

A scrambling method in the frequency domain can be more efficient than that in the spatial domain. There have been, therefore, lots of studies on scrambling in the frequency domain such as DCT (Discrete Cosine Transform), DFT (Discrete Fourier Transform) and DWT (Discrete Wavelet Transform). The scrambling methods in the frequency domain include DCT-based scrambling, wavelet-based scrambling, bit scrambling and scrambling with an encryption algorithm.

2.2 Certified e-Document Authority

As shown in Figure 1, Certified e-Document Authority offers 3 types of primary services for documents, 'storage, transmission/reception and certification', and as additional service, scanning service for digitization of paper documents and web interface for service users, and as service linked with Time Stamping Authority (TSA), time-stamping service for user authentication and verification.

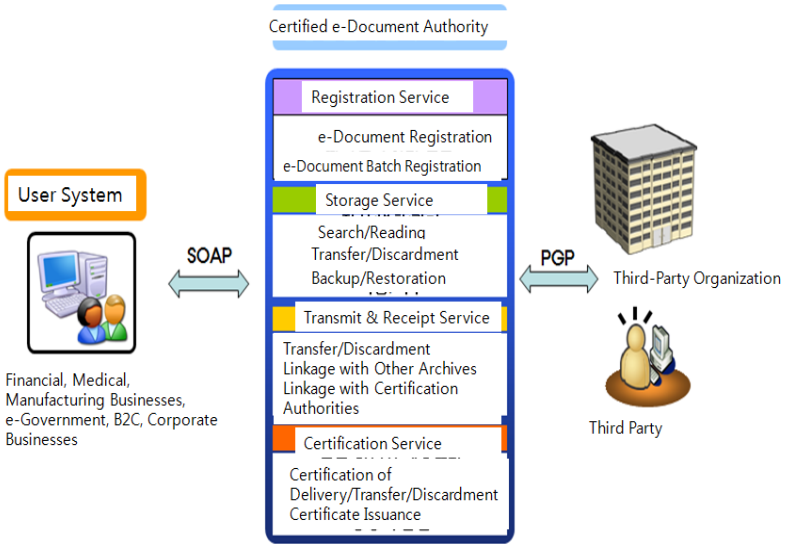


Figure 1. Organization and Service of Certified e-Document Authority

The primary functions of Certified e-Document Authority (document storage and transmission/reception) have direction relations with the management of measures to maintain the authenticity of electronic records (e-records); making out and delivering documents are related with the identity of e-records; such functions as document conversion, management, search, reading, life cycle, transfer/discarding, storage media, backup, restoration, encryption, and prevention of forgery and alteration are directly related with measures to maintain the integrity of e-records; and the function of certification is to certify that an e-record stored in the authority is authentic.

2.3 RFID system

RFID system is composed of a transceiver that reads information of tag, a tag or transponder that provides information, and a back end database that records tag data transceiver collects [4, 6, 8].

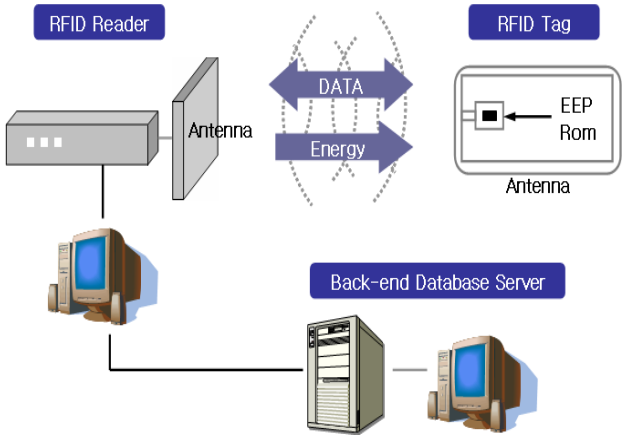


Figure 2. Structure of RFID system

RFID tag is composed of a microchip tag to save data and an antenna to transmit data, and it can be classified into many forms like existence status of chip and frequency. RFID can largely be divided according to the existence of power [1, 2].

-Active tag : It can send radio wave far away because it is connected to the battery which is its power but it is bigger than the passive tag and expensive.

-Passive tag : It is light, inexpensive, and has longer lifespan but its receiving distance is short and quite amount of power must be supplied to transceiver because it is used by receiving energy from transceiver.

The transceiver sends radio wave to tag and the passive tag obtains power by using that radio wave. The passive tag is vitalized with this power and data of tag can be transmitted to the transceiver or can be saved. The active tag periodically transmits signal for various transceivers where data is spreaded like beacon to capture it. The transceiver can be the device that can be carried around or a fixed device installed at the entrance of tollgate. The receiver also has an antenna to send and receive and it is composed of transceiver and processor for decoding data. Transceivers must be operated with one radio frequency usually. If transceivers made from different manufacturers use different frequency, then a lot of expense is required to purchase transceivers of all frequency bands because distributors do not know what frequency has to be used.

3. A Proposed System

The electronic id card forge and falsification prevention method proposed in this study suggests a method to reinforce security of scrambling technology proposes.

The algorithm this study is suggesting first extends the original picture that is used in electronic identification card into hash function using the encrypt key and then rearranges pixel of each image with the created value through scrambling algorithm. This rearranged

image data is inserted into electronic identification card and prevents forge and falsification. The entire structure of the suggesting system is as the Figure 3.

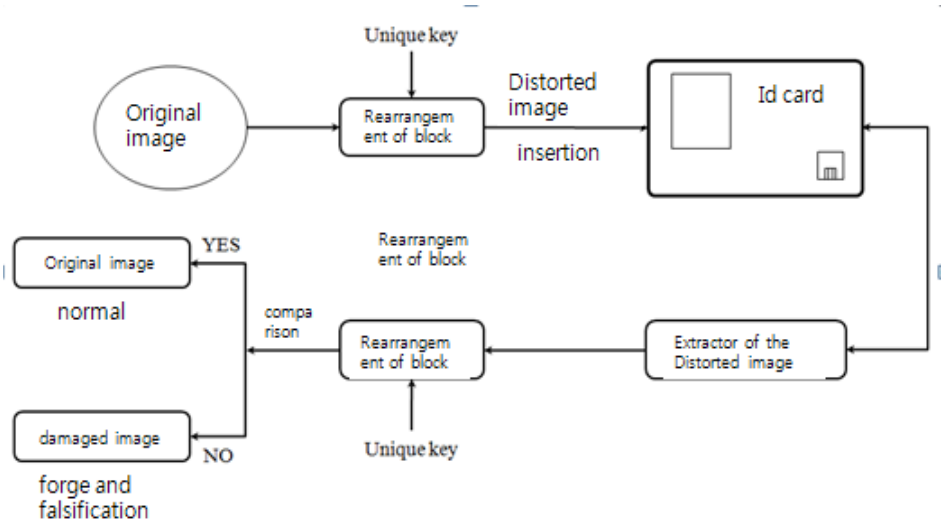


Figure 3. Entire structure of suggesting system

The suggesting system separates the original image using block rearranging algorithm and unique key with scrambling method, rearranges it, and inserts the distorted image to the smart chip of identification card.

Forge or falsification status can be confirmed by extracting the distorted image in the smart chip of identification card using an image extractor and comparing it to the original image using block inverse arrangement algorithm and unique key and check if it matches the original image.

3.1 Scrambling Algorithm

Divide image into a pixel unit before encrypting the image data of electronic identification card. The original RGB color information of the divided image is saved in arrangement. This saved RGB color value is used in encrypting the image.

In order to encrypt each pixel, the encrypt key must be created first. The length of encrypt key is obtained with the method like the following [Formula 1].

$$Key_{length} = (image_{width}) \times (image_{height}) \times RGB_{length} \quad \text{[Formula 1]}$$

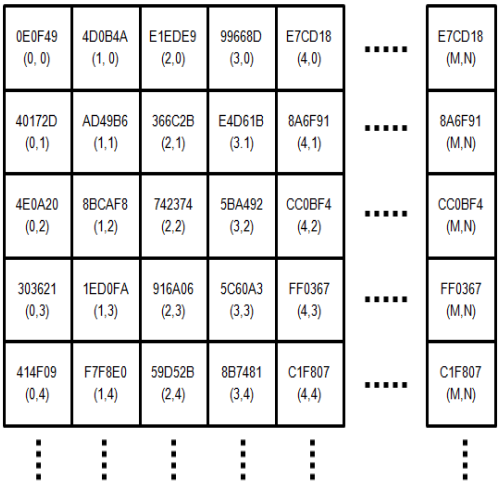


Figure 4. Color Information for Each Pixel

The length of encrypt key creates the encrypt key by multiplying RGB color value to the width and length value of the image. RGB color value multiplies 3, which applies to the length of encrypt key needed, because it uses 3byte(224). In other words, RGB color value of (0, 0) 1 pixel in the Figure 4 is 6 digits and (1,0) also is 6 digits. It is a length of a number that multiplies 6(RGB Length) to the total image pixel value (M × N).

In order to create an encrypt key, repeat encrypt key using hash function (sha-512) until it is same or bigger than the length of encrypt key. Sha-512 function creates 128 letters for value from 0 to F, and the method of creating an encrypt key using hash function for the relevant result value with recursive function is as the Figure 5.

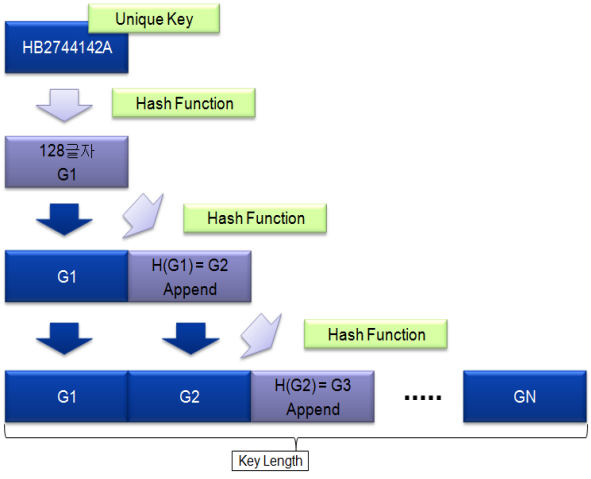


Figure 5. encrypt key creating method

As the result of getting unique value calculating hash function, 128 letters (G1) can be obtained. The method of adding the other 128 letters (G2) with performing the first hash function with the obtained value 128 letters with performing hash function is used..

Original encrypt key = "HB274412A"

encrypt key using hash function =

"D716A4188569B68AB1B6DFAC178E570114CDF0EA3A1CC0E31486C3E41241BC6
A76424E8C37AB26F096FC85EF9886C8CB634187F4FDDFF645FB099F1FF54C6B8C"

encrypt key that used hashing function on the existing result value =

"D716A4188569B68AB1B6DFAC178E570114CDF0EA3A1CC0E31486C3E41241BC6
A76424E8C37AB26F096FC85EF9886C8CB634187F4FDDFF645FB099F1FF54C6B8C342
C3B0163BE0471D6CF48F8CC1DC179D44653A73EDFBB318498B8A8F5A4CEC5138916
09919173A2F4148AAC6DB22DB614CB5EBB47243561FA36F643B2346561"

Like this, length of encrypt key must be more than length of the value that multiplied 6, which is the color length of RGB, and total pixel number.

.

$$EncryptKeyLength \geq \left(\sum_{x=1}^{i}\sum_{y=1}^{i} pixel(x,y) \times RGBLength\right)$$ [Formula 2]

The result value of hash function is printed in hexadecimal, it has a characteristic of RGB color expression method and easy to calculate, Sha512 returns the biggest digit number of hash function as a result, and it was used because of its safety.

The scrambling encryption method for each pixel is same as the [Formula 3] and an algorithm applied to the scrambling method by adding the encrypt key to the color value of the existing original image and then encrypted through color information and calculation about electronic currency saved in arrangement is as follows..

Image_translation_RedPixel(i, j) =

RedPixel(i, j) + HashValue(i*j + j*6, 2))

Image_translation_GreenPixel(i, j) =

GreenPixel(i, j) + HashValue(i*j + j*6 + 2, 2))

Image_translation_BluePixel(i, j) =

BluePixel(i, j) + HashValue(i*j + j*6 + 4, 2)) [Formula 3]

The value results from the modular calculation with 10016(decimal number : 256) after adding hash value to the color information of the original image is information value of the encrypted color. Each red color value, green color value, and blue color value of the number 10016 here is divided into 256 colors. The following Figure 6 is the process of encrypting using the encrypt key and the original color value..

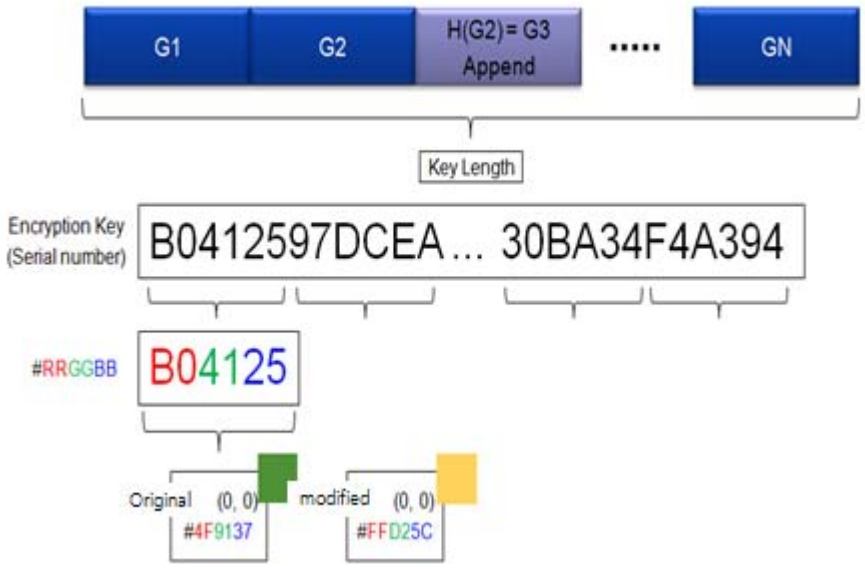


Figure 6. Calculation process of encrypting

From the encrypt key obtained using hash function, it is applied to encryption calculation in 6 digits. For example, it is calculated with RGB color value of the image of first pixel (0,0) using "B04125" 6 digits from the encrypt key. Calculate by substituting red color value "4F", green color value "91", and blue color value "37" from color value "4F9137" of the original image to the [Formula 2].

The result of calculation that applies to (0, 0) pixel about red color value is "(4F16+B016) mod 10016 = FF16", about green color value is "(9116+4116) mod 10016 = D216", and about blue color value is "(3716+2516) mod 10016 = 5C16". By adding this, the color information value of the original image "4F9137" is encrypted into "FFD25C". The encrypt is created with repeated performance of the following algorithm.

This changed image can be restored to the original image when block algorithm and encrypt key are known.

4. Evaluation Of Performance

As shown in Figure 7, the suggested system was so made that the time taken by the system for encryption is tested.

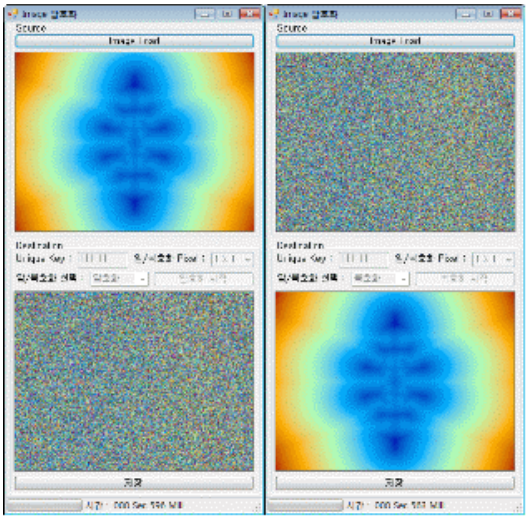
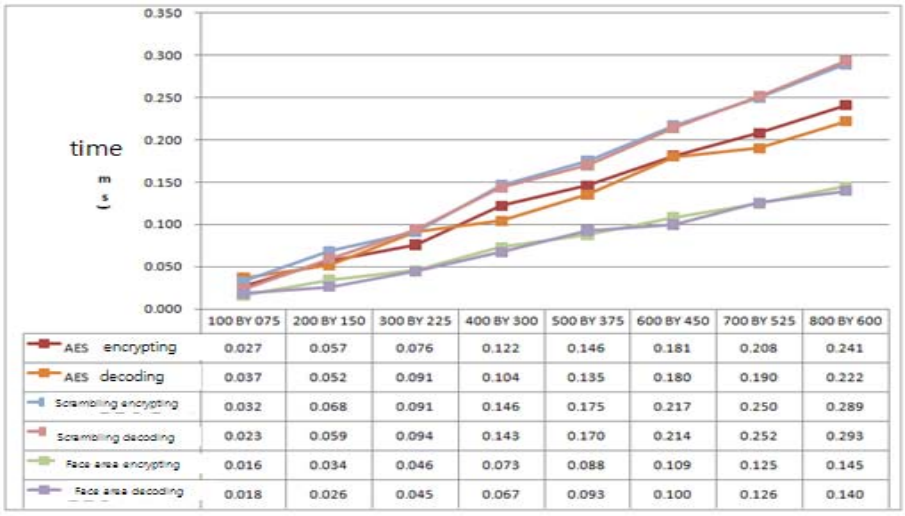


Figure 7. Encryption/Decryption Interface

To evaluate the performance of the suggested system, a performance test was conducted for a method with an AES encryption algorithm, which has been applied in the existing systems, and for a method with the scrambling encryption technique. And, encryption methods and safety for image files were verified per extension (JPEG, BMP, and PNG)

4.1 Comparison in Encryption Speed with a System with AES Algorithm

Encryption methods with an existing AES algorithm and with the proposed scrambling technique were compared in encryption speed as shown in the figure below. The comparison test included image files with different sizes and was conducted in the environment of Quad 2.6GHz. As shown in Figure 8, the scrambling encryption technique is slower than the AES encryption algorithm in case an entire image is encrypted. As a result, the speed of the suggested method becomes improved and better than that of the symmetric-key algorithm in case only a certain part, like a face of a person, is encrypted. The data used in the test was an image including a face, and when about 40% of the face part was included in the encryption, the speed of the suggested method became better by roughly 40% than that of the AES algorithm. Hash algorithm has slower performance speed than AES algorithm but it is simply only the part to create a key, and as other calculation conducts XOR so it obtains faster execution value in the aspect of speed. In case of encrypting the face area, it showed more improved speed as unnecessary parts were not encrypted. Moreover, the speed can be improved in cases of portrait pictures such as ID picture if only parts including information like face are encrypted, not the entire area including background...



	100 BY 075	200 BY 150	300 BY 225	400 BY 300	500 BY 375	600 BY 450	700 BY 525	800 BY 600
AES encrypting	0.027	0.057	0.076	0.122	0.146	0.181	0.208	0.241
AES decoding	0.037	0.052	0.091	0.104	0.135	0.180	0.190	0.222
Scrambling encrypting	0.032	0.068	0.091	0.146	0.175	0.217	0.250	0.289
Scrambling decoding	0.023	0.059	0.094	0.143	0.170	0.214	0.252	0.293
Face area encrypting	0.016	0.034	0.046	0.073	0.088	0.109	0.125	0.145
Face area decoding	0.018	0.026	0.045	0.067	0.093	0.100	0.126	0.140

Figure 8. Comparison of Encryption/Decryption Speeds

This result was fulfilled by comparing it with the method that applied the existing AES encrypting and the scrambling encrypting method is more suitable for low specification devices by adding bit value of image graphic through hash function instead of encrypting method that has many process amount when considering low processing ability and low memory capacity in PDA or other mobile devices that can recognize electronic ID card..

4.2 Safety Evaluation

The current Federal Information Processing Standards (FIPS) published by the National Institute of Standards and Technology (NIST) rate security levels on a scale of 1 to 4. The 128-bit AES algorithm is still assessed to be safe. Such an encryption algorithm, however, has a weakness that it encodes all attribute values and so cannot decide which data is encrypted. On the other hand, the proposed system applies a scrambling technique to an image file and so changes only color values for pixels, keeping the attributes of the image as they are. In case the system changes the values of locations for each of the pixels, instead of changing bit colors, a user can check immediately through a histogram if it is same as the original data. In addition, the system enables the user to correct the encryption algorithm by setting an area for scrambling, in order not to encrypt unnecessary parts.

If a decryption process is conducted without a key, but with an image only, the following result can be assumed: the data used in the test is an image with 100 x 75 pixels in a BMP file with 24-bit RGB colors, so the probability of finding an image with the same number of pixels as the original data is 1/125,829,120,000 (7,500 × 16,777,216). It is, therefore, practically impossible to know which colors are arranged in the original data. Unlike the cases of existing encryption algorithms to check whether an encrypted message is an image file or not, one cannot judge if an image with pixels, where the attributes of the image are left as they are, is encrypted. For example, if a picture with a black dot on the white background is encrypted, it is impossible to know if the original data includes several dots, like a 2-dimensional bar-code, or has no dots on the white background, or has any messages.

Since a hash function is a one-way function, it was impossible to know the original value. And, a method of recursive function call was applied as a stream encryption technique.

5. Conclusion

Use frequency of certificate such as ID card or passport that can prove one's ID is sharply increasing due to development of the 21st IT technology. However, in the existing studies, one's face data that is used in electronic entering system or e-passport is inserted without any encrypting process or through an encrypting process using regular encrypt key. So this study proposed an encrypting system for protecting RFID using a scrambling method.

The scrambling encrypting method was suggested that can be used in low specification mobile devices by changing value about each pixel with relatively fast hash function instead of the complicated encrypting process.

After the suggested system was designed and realized, validity of the suggested system was assessed by safety and calculation speed through performance assessment, and this was developed with purpose of simplifying calculation process in order to apply it to RFID.

As a future study task, more lightweight protocol is needed in certification method for devices with unfeasible calculation ability. In contrast to that composition about each pixel can encode lossless compression in original, a little bit of noise condition was occurred in loss compression like JPEG. This is because it is saved by using similarity level about adjacent color in JPEG compression process and more study about this is needed. An image that official electronic document is saving uses the JPEG2000 format, so if this can be solved then it is expected encrypting as different algorithm of the existing in many other places in the future...

References

[1] K. W. Lee, D. K, Oh, J. Kwan, S. H. Oh, S. J. Kim, and D. H. Won,“Safe Challenge-Response-Based RFID Authentication Protocol, Proper for Distributed Database Environment”, The Journal C of Korea Information Processing Society, Vol. 12-C, Issue No. 03, pp.309~316, 2005.

[2] S. H. Yoo, G. H. Kim, Y. H. Hwang, P. J. Lee, “Stateful RFID Authentication Protocol”, The Journal of Korea Information Processing Society, Vol. 14, Issue No. 6, 2004.

[3] H. S. Jo, "Key Management Mechanism and Performance Improvement Measures for Conditional Access System", The Journal C of Korea Information Processing Society, Vol. 8-C NO. 01 pp. 0075 ~ 0087, Feb. 2001

[4] Auto-ID Center, "860MHz-960MHz Class 1 Radio Frequency Identification Tag Radio Frequency & Logical communication Interface Specification Proposed Recommendation Version 1.0.0", Technical Report MIT-AUTOID-TR-007, Nov. 2002

[5] G. Avoine. Privacy issues in RFID banknote protection schemes. Smart Card Research and Advanced Application - CARDIS, pp. 33-48, Kluwer, 2004

[6] RFID Journal, Michelin Embeds RFID Tag in Tires. Available from http:// www.rfidjournal.com, Jan. 2003.

[7] S. A. Weis, S. e. Sarma, R. L. Rivest, and D. W. Engels, “Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems”, Security in Pervasive Computing 2003, LNCS 2802, pp201-212, Springer-Verlag Heidelberg, 2004.

[8] S. A. Weis, “Security an Privacy in Radio-Frequency Identification Devices ” MS Thesis. MIT. May 2003.

[9] P. Loo, N. Kingsbury, “Watermark detection based on the properties of error control codes,” Vision, Image and Signal Processing, IEEE Proceedings, Vol. 150, Issue. 2, pp.115-121, Apr. 2003,

Authors

Jung-Oh Park

He received the B.S. degree in Computer Science at Sungkyul Univ., Korea, in 2000, and the M.S. degrees in Computer engineering from Myongji Univ., Korea, in 2003. He is currently working towards a Ph.D. in computer science from Soongsil Univ., Korea. His research interests include Network Security, Cryptography and Information Hiding.

Sang-Geun Kim

He received a B.S., M.S. and Ph.D. degree in computer science from ChungAng University, Seoul, Korea in 1987, 1989 and 1996, respectively. Since 1996, he has been a professor in the Division of Computer Engineering, Sungkyul University