2009 International Conference on Availability, Reliability and Security

Making Use of Human Visual Capability to Improve Information Security

Masakatsu Nishigaki[*,**]
[*]Graduate School of Science and Technology,
Shizuoka University
Shizuoka, Japan
[**]Japan Science Technology and Agency, CREST
e-mail: nisigaki@inf.shizuoka.ac.jp

Takumi Yamamoto[*,***]
*Graduate School of Science and Technology,
Shizuoka University
Shizuoka, Japan
[***]Research Fellow of the Japan Society for the
Promotion of Science (DC2)
e-mail: f5745037@ipc.shizuoka.ac.jp

***Abstract*—This paper describes how to make use of human visual capability to improve information security. Here in this paper, two pilot studies are shown; a content protection scheme with image tainting, and a user authentication scheme with unclear images.**

Keywords: security enhancement, human visual capability, copyright protection, image tainting, image-based user authentication, schema

I. Introduction

As everybody knows, there is no perfect security solution on any information system. For instance, even if digital content is protected with data encryption, the purchaser who has its decrypt key can decrypt the encrypted content and distribute the decrypted content (plane content) illegally. Also, we know that security and usability is trade-off. The usability of a system is usually getting worse as user needs more tightened security to the system.

In the authors' opinion, to overcome those kinds of imperfection and trade-off related to security, it is necessary to make use of human psychological and/or cognitive characteristics. This paper shows two pilot studies [1][2] that achieves improvement of the security of a content protection and a user authentication by making use of human visual capability.

II. Image content protection using human visual performance characteristics

A. Illegal copying of image content

Nowadays, digital content is distributed via several media such as Internet and DVD and so on. On the other hand, there is a problem of illegal copying. This paper focuses on image content.

A content provider is able to send image data only to a purchaser by encrypting the image and sending it over the Internet. Only a purchaser who has the decrypt key can obtain the original image by decrypting the encrypted image. However, if a purchaser is malicious, the purchaser can make copies of the decrypted original image and distribute them illegally. Even if image data is provided in a self-decrypting executable file format, a malicious purchaser can capture the original image by hitting the print-screen button on the keyboard, since the original image is on screen when the purchaser runs the self-decrypting executable image. Once a purchaser has captured the original image and then saved it as BMP/JPEG/GIF image file, the purchaser can use it at will. As described above, cryptography [3][4] enables us to protect image when it is distributed, but not when a purchaser uses it.

Watermarking [5] is another technique for image data protection. A content provider embeds watermark in image data and send it to a purchaser. If a purchaser gives image illegally to someone, by checking watermark, a provider can detect who made the illegal copy. Watermarking is expected to be a strong deterrent to illegal copying. However, it is not easy for a content provider to check all images over the Internet. Especially, watermarked image will be never detected as long as a malicious purchaser keeps the illegal copy secretly in local storage of his/her PC. In addition, watermark could be removed by a certain modification or transformation of the watermarked images. Once a malicious purchaser has removed the watermark, the purchaser can use the image at will.

B. Image Tainting and Dynamic Restoration

To solve the problems mentioned above, we propose to distribute a "tainted image" [1], instead of the original images. Here, any image content is intentionally tainted. The tainted image has less quality of image than original one, so that a content provider can distribute the tainted image free. Another important definition of the "image tainting" is that the tainted image dynamically restored on viewer's eye. This means that viewer can perceive the original image, while the tainted images are displayed on viewer's PC.

1) Brightness Modulation

Image tainting by brightness modulation scheme is shown in Fig.1. Firstly, a content provider makes two copies

978-0-7695-3564-7/09 $25.00 © 2009 IEEE
DOI 10.1109/ARES.2009.157

of the original image. Let us call these copies as image 1 and image 2, respectively. Secondly, the content provider increases brightness (RGB color value) of a pixel of image 1 by α, while decreases brightness of the corresponding pixel of image 2 by α. Here, α is positive/negative random value and could vary at each pixel. We call this operation as a "brightness modulation". As a result of the brightness modulation, a content provider obtains tainted images 1 and 2. Finally, the contents provider sends both the tainted images 1 and 2 to a purchaser. The original image is kept secret by the content provider. An example of tainted images is shown in Fig.2, where Fig.2(a) is a original image, and Fig.2(b) and Fig.2(c) are the corresponding tainted images.

2) Dynamic Restoration

The tainted images are dynamically restored on viewer's eye. Figure 3 depicts how "dynamic restoration" of the tainted images takes place. A purchaser runs a dedicated image viewer for dynamic restoration of a set of tainted images. This dedicated viewer displays all the tainted images in turn at the intervals of 1/120 second or shorter. The intervals are so short that the naked eye cannot distinguish them one from another and these images are therefore blended together. The blending occurs at the viewer's retina. Such blending is known as additive color mixture [6][7]. Therefore, α applied for image 1 and -α applied for image 2 are balancing each other out so that a purchaser can perceive image that has a quality equivalent to the original image.

Through the experiments, we have confirmed that we can perceive Fig.2(a) when displaying of Fig.2(b) and Fig.2(c) in tern and fast.

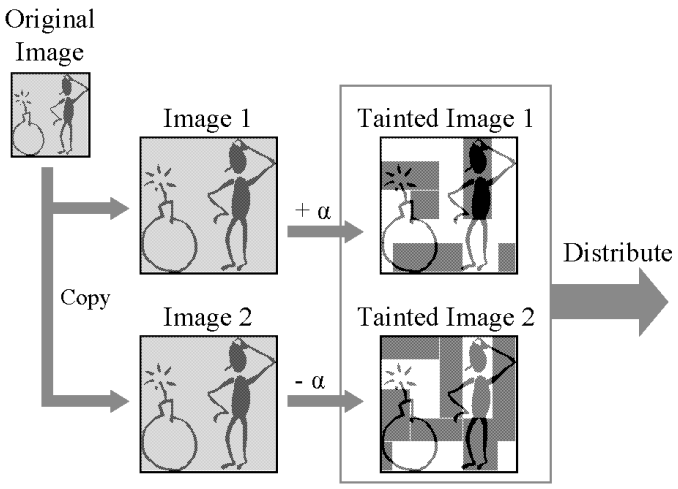

Fig.1. Brightness modulation.

3) Security

In the brightness modulation and dynamic restoration, a purchaser can perceive the original image, while the tainted images are displayed on the purchaser's PC. There exists no original image anywhere. Hence, even if a malicious purchaser hits the print-screen button, what he/she will get is either the tainted image 1 or the tainted image 2.

However, if a malicious purchaser can get both the tainted images 1 and 2, the purchaser will be able to synthesize the original images by calculating the average of these two images pixel by pixel. Alternatively, if a malicious purchaser gives all the tainted images and the dedicated viewer to someone, the someone can perceive the original image, too. That is, prevention of illegal viewing is still not achieved by the brightness modulation and dynamic restoration. However, it is expected that by exploiting some personal characteristic such as an individual's eyesight (the power of vision), the brightness modulation scheme could be cope with illegal viewing.

(a) Original image.

(b) Corresponding tainted image 1.

(c) Corresponding tainted image 2.

Fig.2. An example of tainted images.

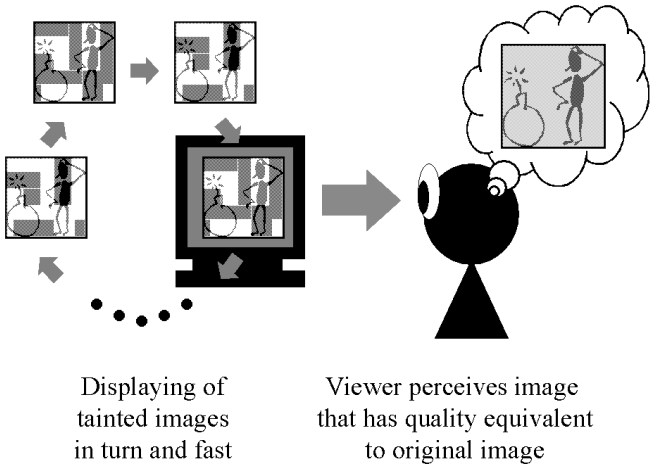

Fig.3. Dynamic restoration.

III. Image-based user authentication using schema of visual memory

A. Shortcomings of password and image-based authentication

Although password-based systems are now widely used in all kinds of authentication, they have some shortcomings in its neglecting of a human limitation. On the password-based systems, if a user chooses a short or meaningful password, it can easily be guessed by a password crack program. To avoid this, users must choose secure passwords (long and random strings). However, most of users prefer to use simple passwords or hesitate to change them frequently since it is not easy for humans to remember a long and random string. In fact, it is known that many users tend to use their names or birthdays as their passwords, to write down their passwords in pocket notebooks, or to reuse the same password in different cases of authentication. These humans' behaviors degrade the security of the authentication system [8][9].

To cope with these shortcomings, image-based user authentication systems using "pass-images" instead of passwords have been studied for reducing the burden of memorizing passwords. The authentication based on recognition of pass-images [8][9][10] is especially effective since humans are significantly more efficient about recognition of previously seen images than precise recall of passwords.

However, on such systems, there is another problem that it is needed to present a user's pass-image on their display at each authentication trial, so they can be vulnerable against an observing attack (shoulder surfing). An observing attack can be a serious problem for image-based authentication systems since the use of the images makes it easier not only for the legitimate users to remember their pass-images, but also for an attacker to peep and remember them.

B. User authentication using unclear images

1) Use of unclear images

To solve the problems mentioned above, we propose an image-based authentication which employs "unclear images" as pass-images [2]. An unclear image is produced from an original meaningful image by image processing such as grayscaling, mosaicing, and noise adding to the spatial frequency domain. The left image in Fig.4 shows an example of the original image. The right image in Fig.4 is the corresponding unclear image obtained by image processing. Although the unclear image has lost a considerable amount of color and resolution, it still holds a certain degree of information about the original image.

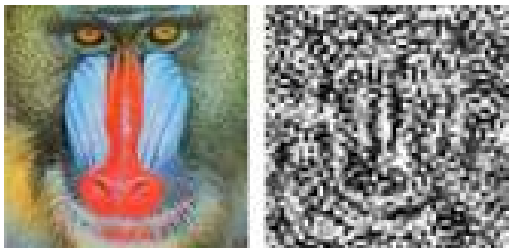

Fig.4. An example of the original image and the corresponding unclear image.

An unclear image still retains some meaning of the original image, but it looks like a meaningless image to users who have never seen the original image before. Even human beings find it hard to memorize meaningless images. That is why, we expected it to be difficult for unauthorized users to memorize legitimate user's unclear pass-images, even if they were allowed to freely observe legitimate users' authentication trials.

The overview of the authentication system with unclear images is shown in Fig.5. Only legitimate users were allowed to see the original images corresponding to their unclear pass-images in the registration phase. By seeing the original images, the legitimate users could recognize the meaning of the unclear pass-images and could easily memorize them by using the original images as cues.

In other words, this scheme gave only legitimate users an underlying knowledge of their unclear pass-images by having seen the original corresponding images. This kind of knowledge is called a "schema" in cognitive psychology [11]. A schema means a structure for knowledge that is unconsciously organized in the human mind when we memorize any incoming information. Once a legitimate user has formed the schema for his/her unclear pass-images that is associated with the original corresponding images, he/she

can easily recognize the meaning of the unclear pass-images.

Therefore, for legitimate users who can memorize unclear images as meaningful images, the burden imposed by their having to memorize unclear pass-images is small, and they will be able to easily find their pass-image among a number of decoy unclear images. On the other hand, it is difficult for unauthorized users to do it. An example of authentication windows is shown in Fig.6.

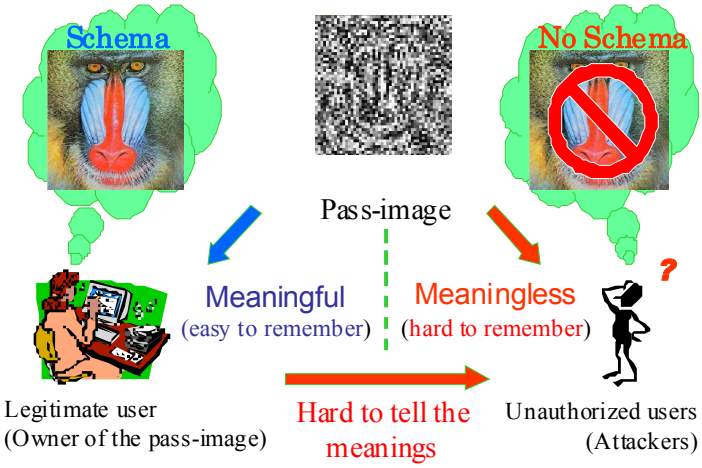

Fig.5. The overview of unclear-image-based authentication system.

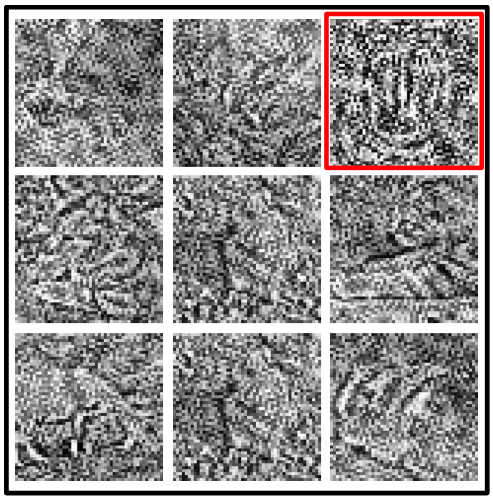

Fig.6. Authentication window.

2) Experimental results

The proposed authentication system has performed well in laboratory-style experiments to quantify the effect of using unclear pass-images compared to the conventional image-based authentication systems which use original image as pass-images. Actually, an experiment with a 9-alternative-type authentication system[1] (Experiment 1) and two kinds of experiments with a 2-alternative-type authentication system[2] (Experiments 2 and 3) were carried out in the literature [2]. In Experiments 1-3, the examinees are 10 male volunteers of college students, and 90 photographs of well-known animals are used as pass-images.

Experiment 1 is to confirm if the legitimate user could remember the unclear pass-images. Firstly, each examinee registers 4 of distinct unclear pass-images. Then, on the following day and 8 days later, every examinee is required to try the authentication. When an examinee can successfully find all the 4 unclear pass-images through 4 rounds of 9-alternative-typed authentication, then the examinee is authenticated. The authentication trial (4 rounds of 9-alternative-typed authentication) is repeated five times with the same set of pass-images. The authentication success rate and the time taken to find out the pass-images among 9 alternatives for each round of authentications are recorded.

Table 1 shows the results of the experiment. All examinees have succeeded in the authentication of 8 days later as well as on the following day. The examinee who has failed once in authentication of 8 days later told us that he had not forgotten any of his pass-images, but just incautiously chosen a wrong image that resembles to one of his pass-images.

Table 1. The results of Experiment 1.

	1 day later	8 days later
Success rate	50/50 (100%)	49/50 (98%)
Average time per round	8.194 sec	7.102 sec

Experiment 2 is to confirm the robustness against an observing attack. In this experiment, the examiner (a legitimate user) chooses the correct pass-image with a mouse click just in front of the examinee (an attacker). The examinee is immediately required to impersonate the legitimate user. On the conventional systems (which use original images as pass-images), the examinees could

[1] In this paper, let us refer to the system in which 1 pass-image is displayed along with k decoy images as "(k+1)-alternative-type authentication".

[2] In this experiment, the number of pass-images that a user should remember is one and the number of the decoy images that are presented along with the pass-image in each authentication phase is one. The system chooses the position of pass-image (left or right) randomly in each authentication trial. That is, the user authentication is completed by choosing one image from two (left or right). The decoy image would be different from each other in each authentication trial.

perfectly remember the legitimate user's pass-image. On the other hand, the proposed system (which uses unclear-images as pass-images) could decrease the attack success rate by about 10 percent.

Experiment 3 is to confirm the robustness against an authentication information leakage with words. The examinee is told the "characteristics[3]" of the pass-image with words. Then, the examinee is immediately required to impersonate the legitimate user. On the conventional system, the examinees could perfectly succeed in impersonating the legitimate user with the information that the legitimate user told them. On the other hand, the proposed system could decrease the attack success rate to 74 percent.

We know that experiments 2 and 3 are considerably advantageous to attackers since the number of decoy images is only one (the 2-alternative-type authentication). Therefore, it would not be an exaggeration to say that both results (decreasing of the attack success rate by about 10 percent and to 74 percent) are big improvements.

IV. Conclusions

This paper introduced two pilot studies to improve information security by making use of human visual capability. In image tainting, image content protection is achieved by using human visual performance characteristics, and in unclear image-based authentication, authentication with a smaller burden of memorizing pass-images and a higher tolerance to observing attack is realized by using schema of visual memory.

Acknowledgments

We thank Dr. Kazuya Shioda of ChanceLab. Corporation, Dr. Atsushi Harada of Mitsubishi Electric Corporation, and Prof. Takeo Isarida of Shizuoka University for their support to this work.

References

[1] K.Shioda, H.Yoshida, M.Soga, A.Takubo, K.Hayashibe, I.Nakamura, T.Mizuno, M.Nishigaki: Digital image content protection using human visual performance characteristics, IPSJ (Information Processing Society of Japan) Journal, vol.46, no.8, pp.2078-2097, 2005. (in Japanese)

[2] A.Harada, T.Isarida, T.Mizuno, M.Nishigaki: A User Authentication System Using Schema of Visual Memory, Proceedings of the Second International Workshop on Biologically Inspired Approaches to Advanced Information Technology, Lecture Notes in Computer Science, vol.3853, pp.338-345, 2006.

[3] J.Zhao and E.Koch: Embedding robust labels into images for copyright protection, Proceedings of ICIPR, 1995.

[4] W.Diffie and M.E.Hellman: New direction in cryptography, IEEE Trans. Information Theory, Vol.IT-22, no.6, pp.644-654, 1976.

[5] R.L.Rivest, A.Shamir and L.Adleman: A method for obtaining digital signatures and public-key cryptosystems, Comm., ACM, Vol.21, No.2, pp-120-126, 1978.

[6] J.C.Maxwell, "Experiments on colour as perceived by the eye, with remarks on colour-blindness", Transactions of the Royal Society of Edinburgh XXI, part 2, pp.275-298, 1855.

[7] J.C.Maxwell, "On the theory of compound colors and the relations of the colors of the spectrum", Philosophical Transaction of the Royal Society of London, 150, pp.57-84, 1860.

[8] Adrian Perring, Dawn Song, "Hash Visualization: a New Technique to improve Real-World Security", International Workshop on Cryptographic Techniques and E-Commerce (CrypTEC), 1999.

[9] Rachna Dhamija, Adrian Perring, "Déjà Vu: A User Study Using Images for Authentication", 9th USENIX Security Symposium, pp.45-58, 2002.

[10] Trevor Pering, Murali Sundar, John Light, and Roy Want: Photographic Authentication through Untrusted Terminals, Trevor Pering, Murali Sundar, John, IEEE Pervasive Computing, Vol 2. No 1, pp.30-36, 2003.

[11] W. F. Brewer, "Schemata." In R.A.Wilson & F.C.Keil (Eds.), MIT Encyclopedia of the Cognitive Sciences, pp.729-730, 1999.

[3] The information that the examiner gave to the examinees is as follows: "the category of the animal (dog, cat, etc.)", "which direction the animal is facing to", "whole body or a part", and "standing or sitting".