Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2006, Article ID 73685, Pages 1–14
DOI 10.1155/WCN/2006/73685

Mutual Image-Based Authentication Framework with JPEG2000 in Wireless Environment

G. Ginesu, D. D. Giusto, and T. Onali

MCLab, Department of Electronic Engineering, University of Cagliari, Cagliari 09123, Italy

Received 30 September 2005; Revised 24 March 2006; Accepted 13 June 2006

Currently, together with the development of wireless connectivity, the need for a reliable and user-friendly authentication system becomes always more important. New applications, as e-commerce or home banking, require a strong level of protection, allowing for verification of legitimate users' identity and enabling the user to distinguish trusted servers from shadow ones. A novel framework for image-based authentication (IBA) is then proposed and evaluated. In order to provide mutual authentication, the proposed method integrates an IBA password technique with a challenge-response scheme based on a shared secret key for image scrambling. The wireless environment is mainly addressed by the proposed system, which tries to overcome the severe constraints on security, data transmission capability, and user friendliness imposed by such environment. In order to achieve such results, the system offers a strong solution for authentication, taking into account usability and avoiding the need for hardware upgrades. Data and application scalability is provided through the JPEG2000 standard and JPIP framework.

1. INTRODUCTION

Nowadays, the deployment of a robust authentication system is one of the most interesting aspects for Internet providers and users. The diffusion of new web services, as e-commerce or home banking, has increased the security vulnerabilities, entailing the need for verifying the identity of both contracting parties and for personal data protection. Against such necessity, the techniques of security breaking are constantly growing together with technology; since attacks become increasingly frequent and well performed. Current auto-cracking tools allow the hackers to gain unauthorized access to digital data, generally with the aim of stealing classified information, as passwords or credit card numbers. In the wireless networks, this problem is still greater as the wardriver community succeed very simply to elude the WEP protocol, traditionally used for WLAN protection. A robust control access system, in addition to privacy and data integrity, becomes the essential condition to support the thriving of World Wide Web and mobile Internet, allowing the identification of legitimate users and avoiding unauthorized intrusion. Furthermore, applications based on a client-server model require to verify the authenticity of service provider, to avoid the risk of coming up against a shadow server.

The most part of current authentication systems is not able to provide these security requirements, especially in wireless environment, where little computational capability, hardware incompatibilities, and poor handiness of user terminals prevent from implementing very complex solutions. For instance, memory-based techniques require the user to precisely recall complex alphanumeric passwords. However, difficulty of password memorizing and poor input interfaces of mobile devices result in the choice of weak passwords, as common words or short PINs, exposing the system to security threats. Besides, these techniques are capable of guaranteeing the identity of user only (*weak* authentication). More advanced solutions have been proposed in order to enforce security and achieve *mutual* or *strong* authentication, that is, the client authenticating itself to a server and that server authenticating itself to the client in such a way that both parties are assured of the others' identity. These methods are based on encryption algorithms, often requiring specialized hardware, as encryption-calculators, tokens, or smart cards. As a result, such solutions are expensive and incompatible with wireless technologies. Consequently, two problems are still to be solved: (i) increasing security and usability of user authentication; (ii) devising a scheme for mutual authentication, possibly for any client's device, from computer terminals to mobile phones. Image-based authentication (IBA) is a valid solution, which guarantees both a high security level without compromising simplicity and efficiency of authentication process. Several experiments of cognitive science show,

in fact, that pictures are easier to recall than alphanumeric passwords [1–3]. Furthermore, graphical passwords do not require hardware upgrades and can be combined with techniques of steganography, watermarking, or image scrambling to insert secret visual information into messages for server authentication.

Several visual login systems have been proposed in the literature, many implementing a weak authentication only. Déjà Vu [4] requires the identification of five random-art images out of a challenge set of twenty-five images. Viskey [5] asks the user to select a series of image spots following a precise order. Picture password [6] and Awase-E [7] require the identification of a correct pass-images sequence, that is, the sequence of images that are chosen by the client during registration, the first employing a single verification stage with a grid of 5×6 images, the second employing multistep stages, each with a number of images depending on the display size. Unfortunately, the process of remembering a combination of abstract images or a precise order of selection may become harder than the use of traditional passwords, thus nullifying the simplification introduced by the visual approach [8]. Furthermore, most of the proposed solutions offer a security level comparable to PIN codes, therefore inadequate to current applications, which require the security of [6–8] character long alphanumeric password. Besides, some of such systems are not suitable for small displays and poor handiness of mobile terminals; Viskey, for instance, may be used only with mouse or light pen. Awase-E, although purposely studied for wireless applications, involves the transmission of a large amount of visual information, which is inconvenient due to bandwidth limitation of wireless channels. GPRS network providers, for instance, generally allow for a bandwidth smaller than 56 kbps, while the billing system is often traffic-dependant. Moreover, all of the above-mentioned IBA frameworks fail in providing mutual authentication. Other graphical systems have been proposed for mutual authentication. For example, a technique of visual cryptography [9, 10] provides each user with a transparency, that is, a portion of visual information, which reveals a secret when combined with another sent by the server during the authentication session. Steganography may be used together with visual cryptography; an overview for such approach is given in [11]. The most widely known technique consists in replacing the last bit of each image pixel with a bit of secret information. These systems rely only on the secret keys exchange; one key is stored into the user terminal, while the other is sent by the server at each login request. So, both the user and the server keys are not very protected against theft or network sniffing attacks, allowing malicious clients or shadow servers to break the security system.

This paper proposes a novel mutual image-based authentication framework (MIBA) that exploits platform scalability in order to achieve a good tradeoff between security and data transfer for several applications and devices, such as computer terminals, PDAs, and mobile phones. While user authentication is implemented through an image-based password creation process, server authentication is granted by the scrambling of any visual information to be transmitted to the client. The proposed framework makes extensive use of the JPEG2000 standard for both image storage and processing, while relying on the properties of wavelet decomposition for the scrambling and transmission of visual information to the client.

The paper is organized as follows: Section 2 describes the wireless connectivity scenario. Section 3 provides a brief overview of the JPEG2000 standard. In Section 4 the proposed IBA method is described in its details. The processes for registration and authentication are illustrated, together with the proposed image scrambling method for mutual authentication and some details related to the JPEG2000 interface. Comparative results are provided in Section 5. Finally, conclusions are drawn.

2. THE WIRELESS ENVIRONMENT

It is recognized that wireless networks are very vulnerable to security issues [12, 13]. Operative systems currently embedded in mobile devices have been implemented in order to optimize the use of available radio resources rather than guarantee an adequate security level. To interfere into a system based on radio-frequency is often very simple.

Three are the basic security requirements defined by IEEE for the WLAN environment, that is, privacy, integrity, and authentication [14]. Privacy ensures that confidential information, as passwords, is not transmitted in clear through the network using cryptographic techniques. Integrity provides that messages are not modified during transmission; it is supported by hashing algorithms. Finally, authentication is needed to verify the clients' identity and to prevent unauthorized access. Many applications also require to authenticate the server: data traffic is only sent after mutual authentication is provided.

Typically, the IEEE 802.11 [14] standard supports the wired equivalent privacy (WEP) protocol to protect wireless communications between clients and access points. It satisfies all security requirements even though with many reserves. In particular, privacy relies on RC4 encryption algorithm and uses a secret key of 64 or 128 bits, which are not sufficient for guaranteeing secure applications. Besides, a simple challenge-response scheme is provided for authenticating only the device; no user and mutual authentications occur.

In order to fix the weaknesses in WEP, a stronger protocol has been recently defined: the IEEE 802.11i [15]. Since it requires hardware and software upgrades, a subset of 802.11i specifications, the Wi-Fi protected access (WPA) has been introduced to offer an intermediate solution, while the whole standard gains acceptance. The main change of 802.11i standard is the adoption of a new encryption algorithm, the advanced encryption standard (AES), which uses 128-, 192, and 256- bit keys. AES is much more robust than RC4, but requires high computational capability for user terminals. For this reason, WPA does not support it and adopts a mechanism still based on RC4, also including a integrity solution. For authentication, IEEE 802.11i can work in two different ways: personal and enterprise modes. The personal mode

performs user authentication through a numeric or alphanumeric password that is stored in the access point and, optionally, also on the user's terminal. It offers a weak level of protection, similar to WEP. The enterprise mode, instead, guarantees for high security performance. It is based on IEEE 802.1X standard [16], requires an external authentication server, and provides for algorithms of mutual authentication.

These protocols achieve security for the wireless portion of connection, between client and access point only. In order to grant end-to-end secure communication and to reinforce wireless security, other types of mechanisms, as end-to-end encryption, password protection, or applications for end-points authentication, must be supplied. For instance, if a user requires Internet access from a wireless network, data protection must be provided on the whole path of communication, together with a mutual authentication system to verify identity of both client and server. The purpose of the proposed approach is then to define an authentication system to provide end-to-end mutual security at application level.

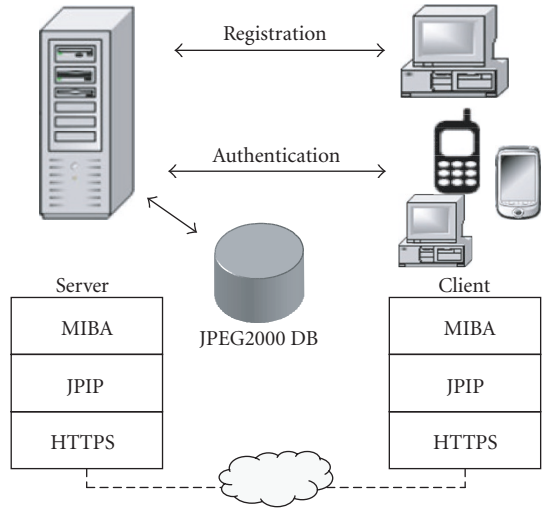

Figure 1: The MIBA framework [21].

3. JPEG2000 STANDARD

JPEG2000 is the state-of-the-art international standard [17–19] for image data coding based on wavelet-domain decomposition and the EBCOT algorithm. The basic system is completely described in its part 1, which gained the status of international ISO standard in 2001. Actually, there exist other 11 official parts, describing several specific aspects of the compression environment.

The basic characteristics exploited in our work are wavelet decomposition and tiling. Decomposition in the wavelet domain is a fundamental aspect of JPEG2000 and is meant to exploit the correlation of visual signal. The image scrambling technique proposed in Section 4.2 exploits the properties of wavelet-domain representation for the introduction of pseudorandom ordering of wavelet coefficients. While JPEG2000 images are generally coded as one block, that is, the whole image is wavelet-transformed and coded as a whole, the standard provides for tiling option. When tiles are used, the coding process is applied separately to each tile, in a similar way to JPEG 8×8 pixel blocks. Although tiling is generally applied to very large images in order to reduce computational complexity, the devised framework adopts tiling as a simple technique for decomposing the images used for authentication and for guaranteeing the scalable transmission of local refinement data.

In addition to the baseline algorithm, our interest is mainly on part 9—JPIP (interactive protocols and API) [20]. JPIP defines syntaxes and methods for the remote interrogation and optional modification of JPEG2000 codestreams and files. It specifies a protocol consisting of a structured series of interactions between a client and a server by means of which image file metadata, structure, and partial or whole image codestreams may be exchanged in a communications efficient manner. For instance, through JPIP the client is allowed to formulate a specific request defining the resolution, size, location, components, layers, and other parameters for the image and imagery-related data to be received. The server responds by delivering imagery-related data with precinct-based streams, tile-based streams, or whole images. Operatively, the JPIP protocol defines how to generate messages out of portions of single JPEG2000 *databins*. Databins contain portions of a JPEG 2000 compressed image representation, such that it is possible to construct a stream that completely represents the information present in a JPEG 2000 file or codestream. For our purpose, JPIP provides for dynamic image data transmission, for example, single regions or incremental refinement information, through client-server interaction.

4. PROPOSED METHOD

The proposed IBA method is based on a client-server interface [21] to optimize processing, minimize data transmission, and improve security. The authentication framework consists of two classical phases: registration and authentication (Figure 1). While registration has to be carried out from a computer terminal, authentication may be performed from any device.

The core algorithm at the base of image authentication consists in an iterative selection and zooming, supported by the JPEG2000 standard, through the use of tiling and JPIP protocol. Such choice allows for data-stream scalability and for an efficient transmission and refinement of image information. Further, end-to-end security is granted by the adoption of the HTTPS protocol, which provides for SSL encryption and, optionally, for authentication. Besides, JPIP allows for scalable transmission of image components.

While scalability, thus data transfer optimization, is assured by the JPEG2000 framework, described in Sections 4.4 and 4.5, mutual authentication is obtained through shared-key image encryption. In fact, during the multistage challenge-response process for authentication, each time the user requests any visual information, the server provides its encrypted version with the key that was defined during the registration phase. The client must then descramble the

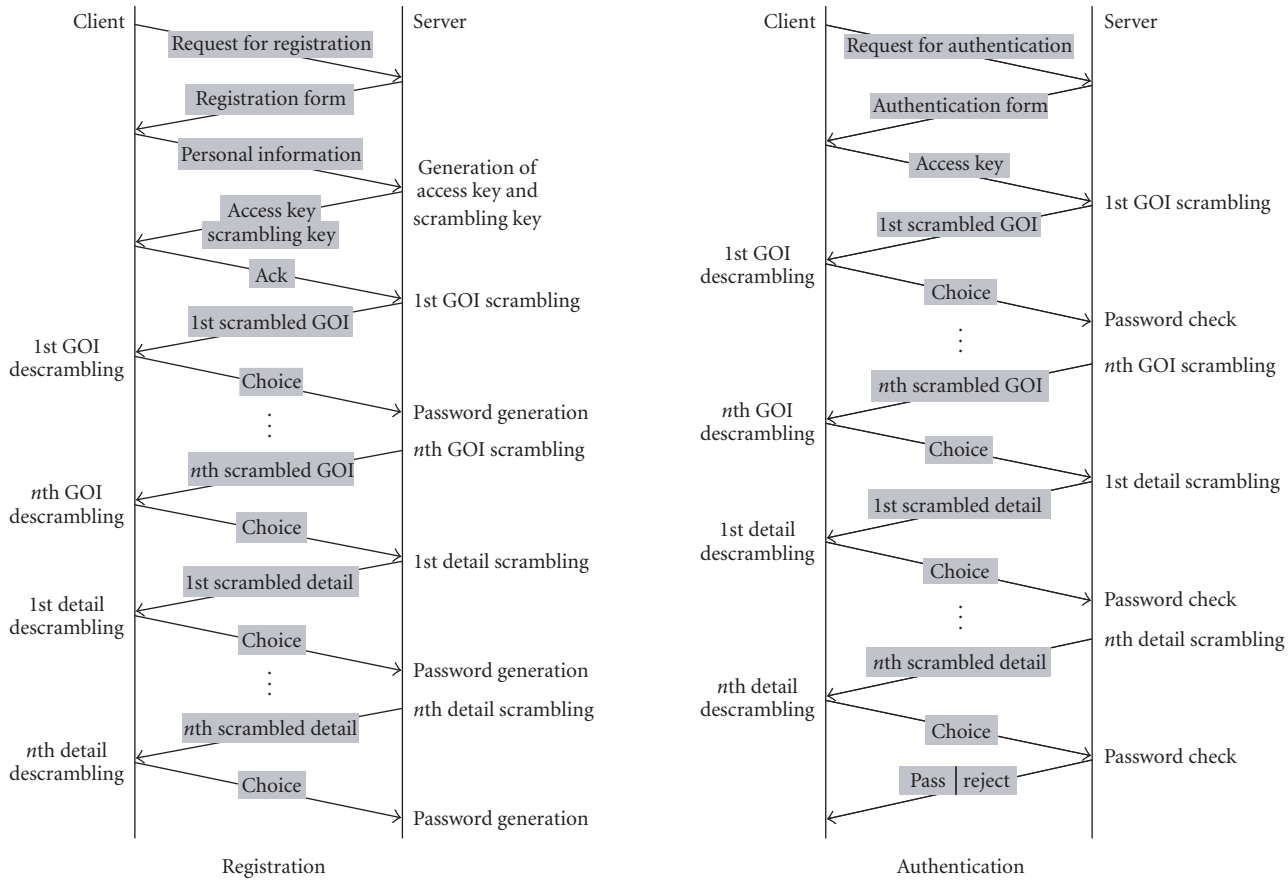

Figure 2: Message exchange scheme for the registration and authentication phases.

visual information in order to make its content understandable. Then there are four possible scenarios.

(1) Trusted server.

(a) *Trusted client*—the transaction may proceed and the scrambling/descrambling process is transparent.

(b) *Malicious client*—the client is unable to understand the visual content. Even if the malicious client gained possession of the scrambling key, authentication would require the visual password identification. Thus, in this scenario the encryption procedure constitutes a double protection against malicious authentication.

(2) Shadow server.

(a) *The server ignores the system architecture*—in this case it will send uncrypted visual information, even though the user always performs the descrambling process. Such process will again result in the encryption of transmitted visual information, thus rendering the image incomprehensible.

(b) *The server knows the system architecture*—the server might try a brute-force attack in order to recreate the correct scrambling key. However, such operation depends in part on the user interaction and the shadow server would have only a few tries. Then, even thou the server succeeded in recreating the scrambling key, it should own the client's pass-images in order to include them among the displayed pictures collection.

In order to minimize data transmission in all environments, the major part of data processing is performed on the server side, which is required to store and manipulate the JPEG2000 compressed images, to generate an appropriate key for the scrambling process, and to perform the image scrambling during each of image authentication. The server replies to each user's request by providing the correct (scrambled) visual information so that refinement data are preferably transmitted. In order to do so, only the correct portion of information, that is, tiles, subbands, and quality layers, is transmitted at each step. On the client's side, the device would only have to perform the descrambling, the exact resizing of the received image, and the transmission of pass-coordinates.

The message exchange scheme for the registration and authentication phases are shown in Figure 2 and will be further described in the following sections.

4.1. Registration

The process of authentication requires the user to define three parameters: an access key, a scrambling key, and the visual password. Such keys have different characteristics and must be defined during the registration process (Figure 2, left). The access key is based on the user's personal data and devices characteristics. It is used to identify the client each time he tries to log in, in order to customize the image-based authentication procedure. Preliminary authentication may be implemented in two different ways through the access key mechanism. While the first consists in defining a shared key to be transmitted each time the user starts an authentication session without intervention, the other requires the user to input some piece of information. Although the second solution is more secure in the case of device theft, the first has been preferred for its simplicity and usability. Then, particular security is not required since the access key has the only purpose of preliminary user identification. Moreover, the case of device theft is generally solved through simple notification by blocking the device or disabling the user's profile (Section 4.6).

The scrambling key is used to generate the pseudorandom sequence that drives the image scrambling process for mutual authentication discussed in Section 4.2. Such key is shared by both server and client, but is transmitted only during the registration phase. Finally, the visual password is generated from the user's graphical choices and is used as authentication password.

Then, the registration interface phase allows the user to acquire his access key, scrambling key, to choose the desired images for authentication and to define the graphical password. During registration, the server first presents a traditional form for submitting the user information. While the access key is directly derived from personal data, the scrambling key is generated through a mixture of personal information and random data, such as the current time or the actual content of a few bytes of RAM. Subsequently, the server shows a large set of images, randomly selected from a database of JPEG2000 images and assembled in GOIs (group of images). These images should be inspired by some different themes, excluding random-art and abstract images in order not to compromise the usability of the proposed method. The user must choose k pass images from the visual database, with the only constraint that one image out of k must be selected only once. For each pass image a single pass detail, that is, the image portion to be used as part of the visual password, must be chosen. Upload of personal images is allowed, although it is generally discouraged, since the authentication process may be easily guessed from personal data. As the registration process may be time consuming and requires the exchange of personal data, it is done online from a computer terminal over secure HTTPS connection.

In order to guarantee data transmission security during registration, HTTPS is adopted with both SSL authentication and encryption. During registration handshake, an SSL secure session is established, including mutual authentication. Then, server and client cooperate in the creation of symmetric keys used for encryption and decryption. In this way, all sensible information, that is, access key, scrambling key, and visual password, are well protected against any form of attack. Such procedure is not adopted during authentication, where only SSL encryption is preserved, while authentication is implemented by the MIBA method itself.

4.2. Image scrambling for mutual authentication

The mutual authentication feature of the devised system is assigned to image data scrambling for the transmission of visual information from server to client. Server's authenticity is then verifiable "at a glance," while the encrypting technique, combined with the visual password, guarantees a higher level of security.

Several image scrambling techniques have been investigated by the recent literature. They are generally based on the randomization of pixels ordering or on the addition of some variations in the coding algorithm. Lossless scrambling/descrambling is defined in [22], using a periodically shift variant (PSV) discrete system in order to permute pixel disposition. Reference [23] performs visual information scrambling through changing the fractional phase in a GF(q^n) composite domain. A method based on chaos system is presented in [24]. It not only permutes the image pixels, but also circularly iterates gray pixel values, through a 2D nonlinear map. Reference [25] discusses two kinds of transformations, based on the Fibonacci and Lucas sequences. They totally decorrelate the visual signal, spreading all pixels, while maintaining equidistance as in the original image, and separating adjacent pixels as much as possible. In [26], the scrambling scheme relies on the 2D extension of the discrete prolate spheroidal sequences (DPSS) is proposed. Other methods define image scrambling in a transform domain. A JPEG-based image encryption algorithm has been proposed in [27]. It consists in three steps: the permutation of luminance and chrominance planes by pseudorandom SFCs (space filling curves); the confusion of DCT coefficients in each DCT block, based on different frequency bands; the encryption of DCT coefficient signs. For JPEG2000 images, scrambling methods are proposed in [28, 29]. Part 8 of JPEG2000 standard, named JPSEC [30], provides for the scrambling to be either performed on the wavelet coefficients or directly on the codestream. Reference [28] presents a system based on JPSEC that encrypts the packet body using RC4 and AES algorithms. In [29], a method for partial-scalable scrambling of JPEG2000 coding units, that is, layers, DWT-levels, subbands, or code-blocks, is proposed. It relies on public-key encryption, which is robust to attacks but results in much more computational cost than secret-key encryption.

Although the previous methods provide several good solutions for the encryption problem, their computational complexity is often high, so that their application may become critical in the case of mobile devices. A choice has been made to develop a simple, yet effective, method, based on the properties of wavelet decomposition. Such choice allows for a nice integration with state-of-the-art coders, such as

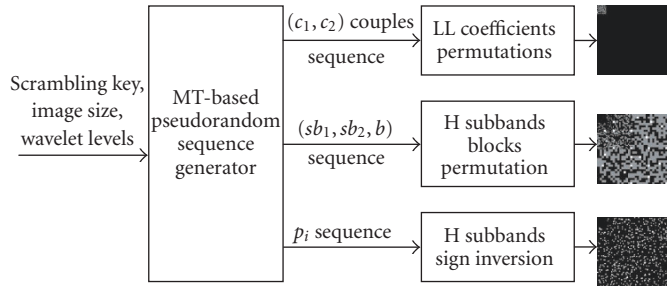

Figure 3: The scrambling method and resulting permutation patterns.

JPEG2000 or SPIHT and adds only an irrelevant computational cost to the codecs. Moreover, the integration of coding and scrambling makes the system more robust to security attacks. As a drawback, the scrambling process inevitably reduces the wavelet ability to decorrelate the signal energy, resulting in weakened coding efficiency. However, such aspect may be restrained so to offer an adequate perceived quality for reasonable compression ratios. In fact, it must be observed that the application of visual authentication is not particularly demanding in terms of visual quality. Thus, the proposed system is based on three stages of pseudorandom permutations in the wavelet domain: LL coefficients, high subbands blocks, and high subbands signs (Figure 3).

The first aspect to be considered is the generation of a pseudorandom sequence of coordinates to drive each of the scrambling stages. The mersenne twister (MT) algorithm [31] has been considered in order to accomplish such task. The method for generating uniform pseudorandom numbers has a large prime period of $2^{19937} - 1$ and consumes a working area of only 624 words and the sequence is 623 distributed to 32-bits accuracy. Since each stage is meant to drive a particular class of coefficient permutations in the wavelet domain, the pseudorandom generator must provide three different sequences from the scrambling key defined during the registration phase. This is obtained by normalizing the MT output to a desired range that covers each permutation's space, depending on image size and decomposition levels. The scrambling key constitutes then the seed for the pseudorandom generator.

While LL coefficients permutation is straightforward, that is, the sequence (c_1, c_2) defines which two coefficients to exchange inside the LL subband, high subband blocks permutation follows a slightly more complex scheme. In fact, the sequence (sb_1, sb_2, b) defines which two subbands sb_1 and sb_2 with indices described in Figure 4 (left), and which reference block b from the largest subband among sb_1 and sb_2 to consider. Block size is proportional to the largest subband size, for example, 2×2 blocks for 32×32 subbands, 4×4 blocks for 64×64 subbands, and so on, so that any subband is divided into 16×16 blocks in the case of square subbands (Figure 4 right).

After determining the largest subbands among sb_1 and sb_2, the reference block position b and block size, the algorithm searches for the block in the smaller subband, which

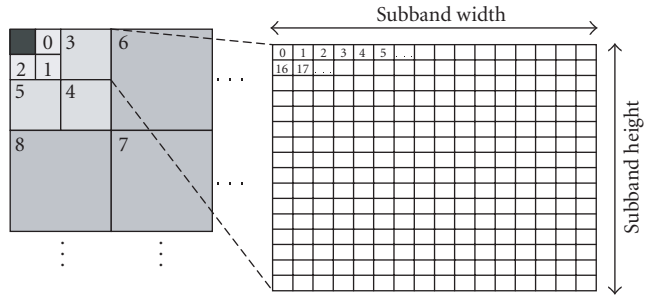

Figure 4: Indexes definition for subband selection (left), and block selection (right).

satisfies the condition of having the least MSE (mean square error) with the reference block (target block). The two blocks of coefficients are then exchanged. Such simple procedure may be schematized as follows:

For each (sb_1, sb_2, b)
  $s_{\max} = \text{MAX}(sb_1, sb_2)$; $s_{\min} = \text{MIN}(sb_1, sb_2)$
  $\text{size}_{\text{reference\_block}} = \text{size}_{\text{target\_block}} = \text{size}_{s_{\max}}/16$
  $\text{position}_{\text{reference\_block}} = b$
  Find target_block in $s_{\min}$ that minimizes
   MSE (reference_block, target_block)
  Permute target_block and reference_block

Finally, sign inversion is driven by the index sequence p_i. Starting from each index, the algorithm searches for the coefficient with greatest absolute value in a neighborhood of

$$\left(\frac{\text{subband\_width}}{16}\right) \times \left(\frac{\text{subband\_height}}{16}\right) \quad (1)$$

coefficients. The sign of such coefficient is then inverted. Both H blocks permutation and sign inversion stages are implemented as a reasonable tradeoff between computational complexity, which is maintained very low, and minimization of the effect of scrambling on compression performance. In fact, the choice to permute blocks with minimum MSE distance and to invert the sign of locally maximum coefficients guarantees that the decomposed signal decorrelation is not dramatically reduced. Another interesting aspect of the

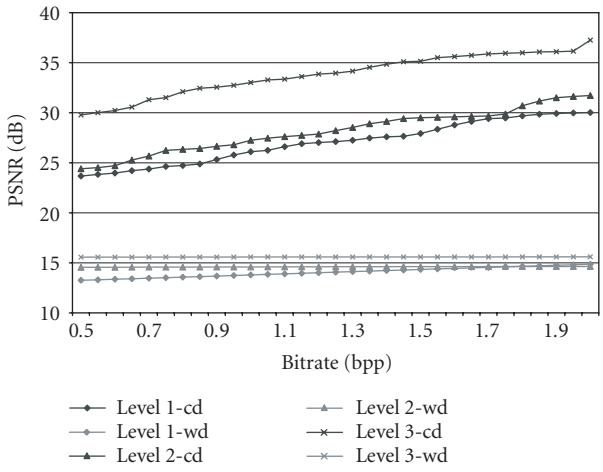

Figure 5: Average coding results for three detail levels with correct (cd) or wrong/no (wd) descrambling.

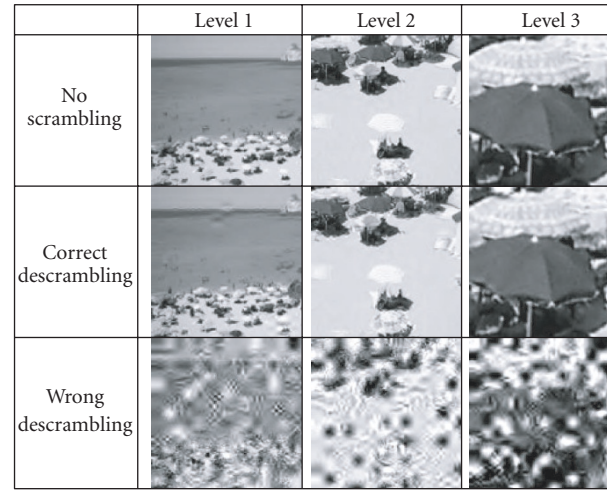

Figure 6: Example of visual results for the scrambling technique, coded at 1.5 bpp.

proposed method is that the descrambling process simply follows the scrambling procedure by reversing the order of each permutation sequence.

In order to evaluate the proposed algorithm in the application environment, 10 different test images have been considered, with three levels of detail each. In Figure 5, the average rate-distortion curve is shown for each detail level, considering correct scrambling/descrambling (cd) and wrong or no descrambling (wd). As expected, higher detail level corresponds to more efficient compression, since the image content decreases accordingly. Moreover, although the scrambling/descrambling process has still an important effect on coding efficiency, that is, there is an average deterioration of 5 to 8 dB compared to unscrambled coding, at a bitrate of 1.5 bpp the system offers adequate image reproduction. This is also illustrated by Figure 6, where a visual comparison between unscrambled, correctly descrambled, and wrongly descrambled images is provided. It must also be observed that wrong or no descrambling, or equivalently wrong or no scrambling with correct descrambling, results in unintelligible image data, achieving a constant PSNR of about 15 dB.

To evaluate computational cost, 10 different test images have been processed with complete codecoding and scrambling-descrambling phases. Compression has been carried out at 16 different rates, ranging from 0.5 to 2 bpp, in order to evaluate the incidence of the proposed scrambling technique with several codec settings. Average results are presented in Figure 7 as the ratio between scrambling-descrambling time and complete processing time. Three different scrambling profiles were used and are reported as L, H, and S, meaning the number of low, high frequencies, and sign permutations, respectively. It must be observed that results shown in Figures 5 and 6 were obtained with the profile $L, H, S = 80\,400\,1000$. As expected, computational cost is inversely proportional to the scrambling profile and decreases for increasing compression rates. With the chosen profile (80 400 1000), the incidence of the scrambling technique is maintained around 10–13% without any code optimization.

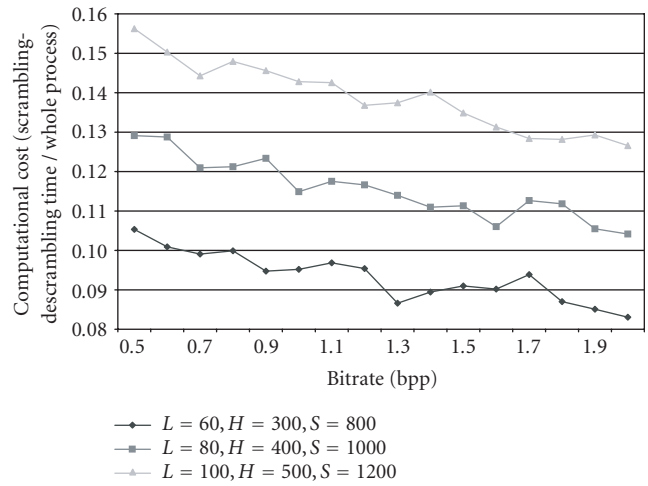

Figure 7: Computational cost evaluation.

4.3. Authentication architecture

The proposed method consists in a challenge-response scheme, which achieves multiple levels of security for both server and user authentication. On the one hand, image scrambling, as described in Section 4.2, provides mutual authentication based on a shared secret key; the server is recognized as trusted only if it owns the user pass images, implements the correct system architecture, and knows the scrambling key. Besides, only a trusted user, which has acquired the access and scrambling keys during registration, may login and decrypt the transmitted images to select its visual password. On the other hand, the IBA architecture guarantees a stronger user authentication, essential in order to avoid

Table 1: Application profiles.

Profile	Device	Connection	Security	(k, h, N)
Low	Mobile	GPRS	Limited	$(1, 9, 9)$
Medium	PDA	Wireless	High	$(4, 16, 16)$
High	PC	LAN	Very High	$(4, 25, 75)$

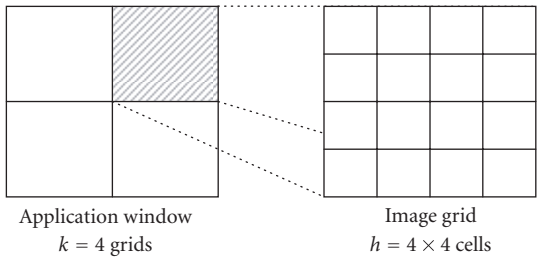

Figure 8: Example of partitioning of the application window.

counterfeit clients' access to the system for stealing private information.

The IBA password consists in the recognition of the pass images and pass details. Device/complexity scalability is achieved through parameterization of this procedure. The application window is divided into k grids, each made of h cells (Figure 8). During the pass image/s selection procedure the user has to correctly identify the k pass image/s among N images, randomly extracted from the JPEG2000 database. Similarly, during the detail selection one secret detail must be recognized for each pass image through the iterative zooming process. By defining with d_{img} and d_{dsp} the sizes of original image and display and the number of iterations for the pass image selection P_1 and for the detail selection P_2 result

$$P_1 \leq \frac{N}{h}, \qquad P_2 \leq \left\lfloor \log_h \left(k \cdot \frac{d_{\text{img}}}{d_{\text{dsp}}} \right) - 1 \right\rfloor. \tag{2}$$

So that the maximum number of iterations is

$$P_{\max} = \max \left[P_1 + P_2 \right]. \tag{3}$$

By choosing a combination of $\{k, h, N\}$, the proposed framework may be easily adapted to any user device. Three application profiles have been defined in Table 1.

4.4. User authentication

During the authentication phase, the server manages the preliminary user and user's device identification by detecting and decrypting the access key. If this is a valid key, the challenge-response scheme based on the scrambling key may start. For each authentication session, the server must send a number of scrambled image sequences between $1 + P_2$ and $N/h + P_2$. Only if the user owns the scrambling key, the received images can be correctly decrypted and displayed. The visual password codes are transmitted step by step, minimizing the risk of sniffing. Whenever the server detects an authentication failure, the authentication process is not interrupted until the last step. Only then, the user is rejected and a notification policy is adopted. During authentication, the user must recognize the combination of k pass images with their pass details. During each authentication session, the server shows k grids, each containing h images randomly positioned in order to minimize the risk of back-shoulder attack. Such randomization does not undermine the method's usability, since the pass image recognition process is not based on image location. After the first stage of verification, the k grids are used to divide the selected images each into h regions. For each image, the user must iteratively select the portion containing its pass detail.

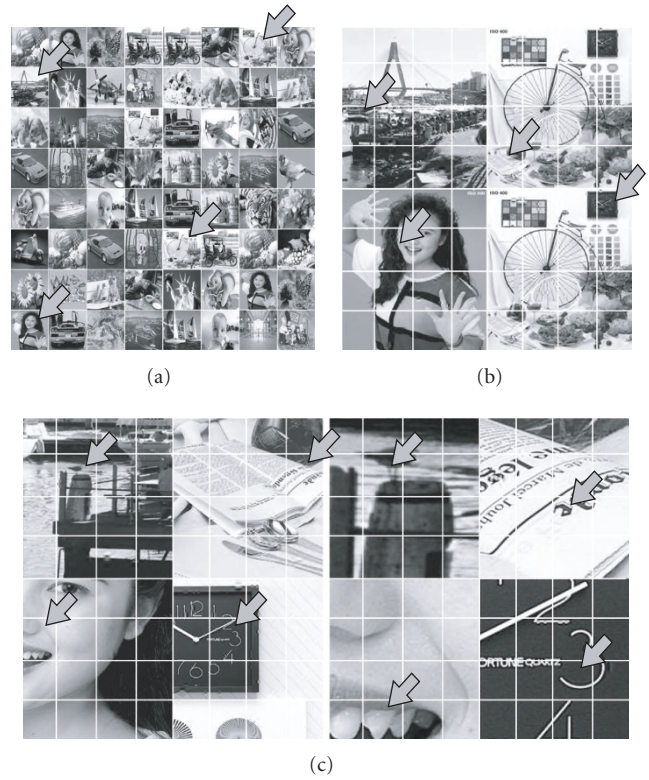

Figure 9: Example of authentication process for the medium profile.

The values of k and h depend on the desired degree of security. As described in Section 4.3, a good tradeoff between security and usability for the medium profile is to use $k = 4$, $h = 16$. An example of authentication is provided in Figure 9 for the medium profile. The time sequence of four authentication steps is shown from 1 (upper left) to 4 (lower right). While step 1 consists in the choice of four pass images (one duplicated) out of 16, the other steps are the recursive pass detail selections. Arrows indicate the user's choice.

Since the proposed framework is devised to work in wired and wireless environments, it is essential to consider the severe constraints on user friendliness and data transmission capability imposed by mobile devices and GPRS technology. The medium profile was conceived for use with PDAs and wireless connection. Nowadays, such devices

offer generous displays and good interactivity, so that decreasing the value of $[h, N]$ to $[16, 16]$ is sufficient to achieve a good tradeoff between usability and security performance. On the other hand, mobile devices with limited connectivity and interactivity require the extreme downscaling of the proposed method. For such reason, the low profile has been set to $k = 1$, $h = 9$, and $N = 9$. In mobile environment, personal device/card codes as the international mobile equipment identity (IMEI) and the subscriber identity module (SIM) may be used to allow for the unique identification of the user every time he logs on the network.

4.5. JPEG2000 parameters

JPEG2000 and JPIP are used in order to transmit only those portions of the scalable image datastream that are required at the client's side at each step. In the proposed method, tile databins are the basic elements of JPEG2000 images used by JPIP. JPEG2000 images are partitioned into 40×40 pixel tiles, coded with 5 decomposition levels and 6 quality layers (0.15, 0.3, 0.5, 0.75, 1.0, 1.5 bpp). Scalability is obtained through the combination of three parameters: tiles, reduce factor (resolution scalability), and quality layers. The number of tiles to be transmitted at each step is proportional to

$$N_{\text{tiles}} = \frac{d_{\text{img}}}{(h^{P-P_1} \cdot d_{\text{tiles}})}. \tag{4}$$

By defining the resizing factor between physical and displayed image portion as

$$Z = \frac{k}{h^{P-P_1-1}} \cdot \frac{d_{\text{img}}}{d_{\text{dsp}}}, \tag{5}$$

the reduce factor may be made proportional to

$$\text{reduce} = \lfloor \sqrt{Z} \rfloor, \tag{6}$$

while the quality layer is assigned the value

$$Q = -5 \cdot \frac{\lfloor \sqrt{Z} \rfloor}{\lfloor \sqrt{Z} \rfloor_{\max} + 6}, \tag{7}$$

where $\lfloor \sqrt{Z} \rfloor_{\max}$ represents the maximum resizing factor with the given d_{img}, d_{dsp}, and $P_{\max}$ values.

4.6. Notification policies

The proposed MIBA method is supported by event-management and notification policies to increase the protection level against unauthorized intrusions. These policies allow legitimate users to control and check all events related to the authentication process, in order to avoid malicious users from registering under an assumed name or accessing through password guessing.

As soon as the registration phase is done, the server sends to the user a confirmation e-mail. The e-mail contains personal data which can be checked to ascertain registration accuracy. Neither authentication keys nor registered images and password are enclosed; in fact, the former should have been already sent through SSL secure connection, while the latter are never transmitted. The e-mail also indicates a URL corresponding to a web page always updated with all the authentication events log. The user may check this page in order to detect immediately any attempt of unauthorized access. Notification is also adopted in case a wrong password is entered. During authentication, errors in password inputting may occur because a legitimate user does not remind its password correctly or a malicious user tries to guess it. In both cases, the server allows up to three attempts. After that, the system is temporarily inhibited and a notification e-mail is sent to the legitimate user, who may modify its password or simply reactivate the system in case of mistake. Such policies constitute a further protection against password-guessing attacks. It must be noted that the notification policies may be set differently, depending on the security level required by each application.

Another notification mechanism is the possibility of physically blocking the mobile device when lost or stolen. By gaining possession of a personal device where both the access and scrambling keys are stored, a malicious individual would be able to try an educated guess attack. To prevent such risk, the stolen or lost device can be physically blocked, for example, mobile phones are identified through the IMEI that is also used to freeze the device permanently. Further, in case of device theft or loss, the legitimate user may inhibit or reset his authentication profile.

5. RESULTS

The proposed method has been evaluated in the medium profile (PDA environment), estimating performance in terms of security, as possible input combinations, data transfer, and usability, as the amount of information required for visual password memorization. Section 5.1 summarizes all authentication scenarios and analyzes possible attacks. Section 5.2 provides a consistent performance comparison between the proposed method and the other visual password techniques. For this purpose, image scrambling is not considered and the analysis is performed in terms of input combinations, data transfer, and user friendliness. Finally, Section 5.3 presents overall results by considering the complete framework.

5.1. Risk assessment

In order to analyze all possible use cases and relative risks, let us first introduce some basic notation. Let us call M the generic malicious entity and use the pedices c, s, and t to indicate client, server, or third party, respectively. An apex with incremental numbering is used to indicate one particular attack occurence, so that M_c^3, for instance, specifies the third case of attack carried out by a malicious client. Similarly we call K the generic key information and use pedices a, s and v to indicate the access, scrambling, and visual key, respectively. Since the visual key is provided through several steps a further numbering is used, for example, K_{v2} indicates the second part of the visual key. The analysis of possible scenarios is split into two main cathegories: (i) either the malicious

Table 2: Classification and characteristics of third party attacks.

Event	Phase	What is stolen	Attack	Likelihood	Impact	
					Value	Notes
M_t^0	—	User device (K_a and K_s)	device theft	Medium	Low	In case of theft, the device/account can be blocked
M_t^1	Registration	Personal user information	Eavesdropping man in the middle	Very Low	Low	K_a is derived from personal information and other data
M_t^2		K_a and/or K_s			Medium/high	Preliminary identification and scrambling/descrambling would be possible
M_t^3		One or more K_{vi}			Low	The value of the visual key is generated dynamically and changes continuously
M_t^4	Registration/ authentication	One or more pieces of scrambled information	Eavesdropping man in the middle	Low	Medium	The visual information is visual useless without the scrambling key
M_t^5	Authentication	K_a			Low	Preliminary identification would be possible
M_t^6		One or more K_{vi}			Low	See M_t^3
M_t^7		The look of one or more K_{vi}	Backshoulder/social engineering	Medium	Low	All other keys should be known

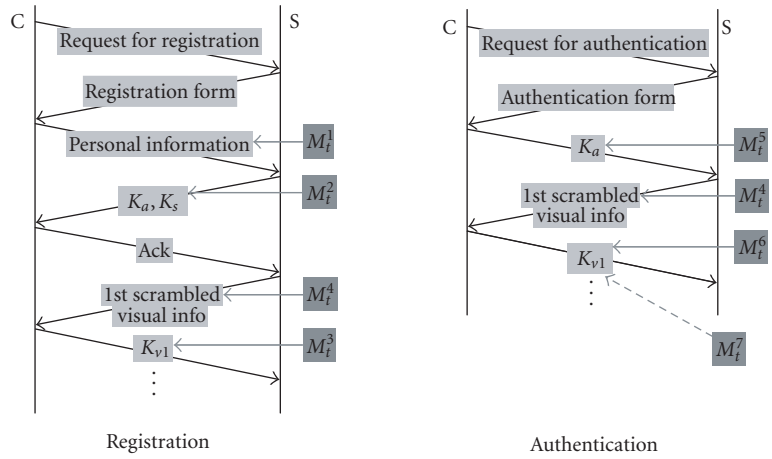

Figure 10: Message exchange and third party attacks.

entity is a third party who tries to acquire sensible credentials during normal client-server interaction (interception), or (ii) attacks are performed by a malicious entity pretending to be the client/server (impersonation or brute force attack).

In the case of third party attack, the malicious entity generally tries to acquire some piece of personal information by managing to break into the client-server transaction. Figure 10 schematizes the authentication and registration processes and pinpoints all possible attacks. In Table 2, third party attacks are summarized and analyzed in order to evaluate their likelihood and impact on system security. A *very low* to *high* empirical scale is adopted.

Attacks performed by malicious clients or through shadow servers generally fall in the category of impersonation attacks (Table 3). The malicious client will try to perform authentication through brute force or educated guess attacks. On the other hand, clients may unknowingly connect to a shadow server and divulge sensitive credentials such as authentication credentials. Both cases require the knowledge of some piece of user information. Evidently, attack likelihood is inversely proportional to the system knowledge.

It can be noted that whenever the attack presents a high impact, its likelihood is low. Security is further discussed in the following sections, while notification policies discussed

Table 3: Classification and characteristics of malicious clients or shadow servers attacks.

Event	What information is known	How was acquired	Possible attack	Impact	
				Value	Notes
M_c^1	Nothing	—	Brute force	Very low	—
M_c^2	K_a	M_t^2 or M_t^4	—	Low	—
M_c^3	K_a and K_s	M_t^0 or M_t^2	Brute force/ educated guess	Medium/ high	—
M_c^4	K_a, K_s, and the look of one or more K_{vi}	M_c^3 and M_t^7	Educated guess	High	—
M_s^1	System architecture	System knowledge	Masquerade	Low	Extremely improbable; the shadow server should have knowledge of the image database and of each user's profile
M_s^2	System architecture and K_s	M_s^1 and M_t^0 or M_t^2		Low/ medium	

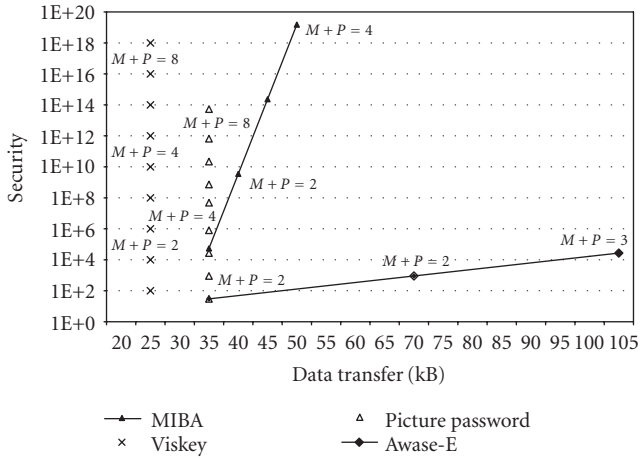

Figure 11: Security against data transfer performance for the medium profile.

in Section 4.6 constitute additional countermeasures against several attack scenarios.

5.2. Framework evaluation

For the medium profile, the performance of the proposed IBA method (MIBA) has been compared with three-state-of-the art graphical password systems compatible with mobile platforms: Viskey, picture password, and Awase-E. Security is reported against data transfer in Figure 11. For the proposed method, security is given by

$$S_{\mathrm{MIBA}} = \begin{cases} N \cdot (N-1)^{k-1}, & 1 \le P \le P_1, \\ N \cdot (N-1)^{k-1} \cdot h^{k(P-P_1)}, & P > P_1, \end{cases} \tag{8}$$

with a limitation on the maximum number of zooming stages P depending on the original image and display sizes, described in (3). In the figure, $M + P$ indicates the length of the password, where M corresponds to N/h in the proposed method, that is, the number of image sequences shown by the server for the pass-image selection. For the other methods, $M + P$ corresponds to the number of images or spots to be recalled and selected by the user. The first authentication step consists in the transmission of composite images and requires 35 KB on average. At each successive step, the size of the JPEG2000 stream decreases progressively thanks to the possibility of refining image information. As a result, an average saturation of the transmitted stream is recorded. On the other hand, Awase-E requires one image of 35 KB on average to be transmitted at each step. Picture password only requires the transmission of one composite image of about 35 KB, whereas Viskey only requires one single image of about 25 KB. These last two methods are then the optimal choice for data transmission. However, their solution is unacceptable because of reduced usability. In fact, the data transfer gain is compromised by the need for choosing an exact combination of images or a precise spot sequence in a specific temporal order. Furthermore, if we consider the security increment given by the scrambling process, the proposed method provides a better protection, allowing for an adequate security despite lower $M + P$ value.

Finally, in order to evaluate the usability, the amount of information that the user is required to recall has been considered. Figure 12 shows the 3D distribution of the considered features: security, data transfer, and usability, in terms of mnemonic load. Mnemonic load is measured as the number of pass images or pass details to be recalled for completing authentication. A multiplicative weight of 2 is considered each time the visual method requires a precise ordering of the pass-image/detail sequence. The triangle and circle marks represent the best and worst situations, respectively.

The proposed method results simpler than all other visual login systems; it only requires the memorization of four pass images and four secret details, independently of data transfer and security level. Awase-E, instead, asks the user to remember one image for each verification stage, at the expense of data transfer. For the same mnemonic load, Awase-E requires eight verification stages ($M + P = 8$), corresponding to the transmission of eight image sequences. Viskey and picture password require to recall a variable number of spots or images, depending on the password length. Moreover,

Table 4: Overall results.

Key length (bits)	Security	MIBA	
		Visual password	Visual password and scrambling
16	65536	(16, 1, 0), (32, 2, 0)	—
32	4.295E+09	(16, 2, 0), (32, 3, 0)	(16, 1, 16), (32, 2, 12)
64	1.845E+19	(16, 4, 0), (32, 5, 0)	(16, 2, 32), (16, 3, 16), (32, 4, 12)
128	3.403E+38	(16, 8, 0)	(16, 2, 96), (16, 3, 80), (16, 4, 64), (32, 4, 76)
256	1.158E+77	—	(16, 2, 225), (16, 3, 209), (16, 4, 193), (32, 4, 204), (32, 5, 188)
512	1.34E+154	—	(16, 3, 464), (16, 4, 448), (16, 5, 432), (32, 4, 460), (32, 5, 444)

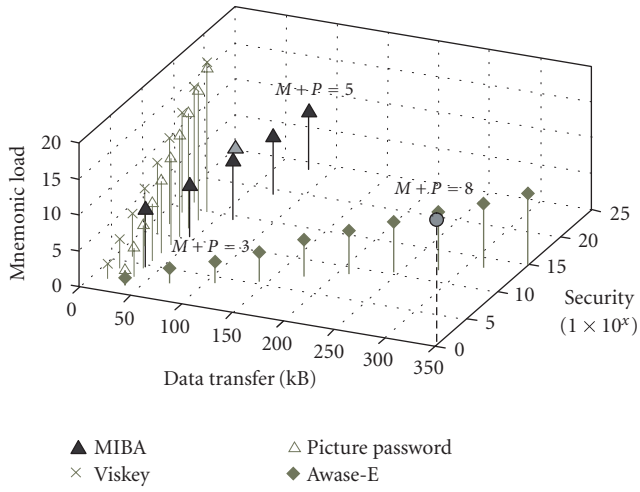

Figure 12: 3D distribution of security, data transfer and mnemonic load for several IBA methods.

a precise selection order must be followed, considerably compromising the system usability.

5.3. Overall results

The security level of the proposed MIBA method is evaluated and compared to a generic system based on a K-bit key. Several MIBA setups are reported in Table 4, achieving the same level of security as the corresponding key length value. Both cases with and without scrambling are considered and represented by the triplet $\{N, P, L\}$, combinations of image alphabet size, number of steps to select the visual password, and length of the scrambling key, respectively. While the visual password alone cannot offer a security level greater than a 128- bit key, the scrambling method allows for a security level comparable to that of any key. Results with scrambling represent the overall security of the MIBA system, excluding the access key input.

CONCLUSIONS

A novel mutual image-based authentication framework has been presented. It consists in a challenge-response scheme based on visual password and image scrambling. This architecture offers strong protection against malicious clients, who might penetrate the system only by taking over both visual password and scrambling key. The risk of impersonation attack by a shadow server is equally unlikely, since the images needed for authentication are transmitted after scrambling. Then, only if the pass images, visual password and scrambling key are successfully stolen on the server side, a malicious entity may impersonate the trusted server. The proposed system may be implemented in any environment by upgrading the user's device with simple software: complexity is minimized in order to be compatible with the limited computational capabilities of some user terminals, as mobile phones. System usability has been taken into account by considering both difficulty of memorization and restrictions of user interfaces, especially in wireless environment. The proposed approach offers a modular architecture and exploits the properties of JPEG2000 and JPIP to achieve datastream and application scalability. Results indicate the validity of the devised method, which realizes the better tradeoff between security, data transfer, and usability in several application environments.

REFERENCES

[1] A. Paivio, T. B. Rogers, and P. C. Smythe, "Why are pictures easier to recall than words?" *Psychonomic Science*, vol. 11, no. 4, pp. 137–138, 1968.

[2] R. N. Shepard, "Recognition memory for words, sentences, and pictures," *Journal of Verbal Learning and Verbal Behavior*, vol. 6, pp. 156–163, 1967.

[3] D. Weinshall and S. Kirkpatrick, "Passwords you'll never forget, but can't recall," in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '04)*, pp. 1399–1402, Vienna, Austria, April 2004.

[4] R. Dhamija and A. Perrig, "Déjà Vu: a user study using images for authentication," in *Proceedings of the 9th Usenix Security Symposium*, pp. 45–58, Denver, Colo, USA, August 2000.

[5] Software and Solutions from Cologne, http://www.viskey.com.

[6] W. Jansen, S. Gavrila, V. Korolev, R. Ayers, and R. Swanstrom, "Picture password: a visual login technique for mobile devices," Tech. Rep. IR 7030, National Institute of Standards and Technology, Gaithersburg, Md, USA, July 2003.

[7] T. Takada and H. Koike, "Awase-E: image-based authentication for mobile phones using user's favorite images," in *Proceedings of the 5th International Symposium on Human Computer Interaction with Mobile Devices and Services*, pp. 347–351, Springer, Udine, Italy, September 2003.

[8] D. M. Wegner, F. Quillian, and C. E. Houston, "Memories out of order: thought suppression and the disturbance of sequence memory," *Journal of Personality and Social Psychology*, vol. 71, no. 4, pp. 680–691, 1996.
[9] M. Naor and B. Pinkas, "Visual authentication and identification," in *Advances in Cryptology (Crypto '97)*, B. Kaliski, Ed., pp. 322–336, Springer, Berlin, Germany, 1997.
[10] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology (EuroCrypt '94)*, A. De Santis, Ed., pp. 1–12, Springer, Berlin, Germany, 1995.
[11] M. Kharrazi, H. T. Sencar, and N. Memon, *Image Steganography: Concepts and Practice*, Lecture Note Series, Institute for Mathematical Sciences, National University of Singapore, Singapore, Republic of Singapore, 2004.
[12] F. Majstor, "WLAN security threats & solutions," in *Proceedings of the 28th Annual IEEE International Conference on Local Computer Networks*, p. 650, Brussels, Belgium, October 2003.
[13] W. Shunman, T. Ran, W. Yue, and Z. Ji, "WLAN and it's security problems," in *Proceedings of the 4th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT '03)*, pp. 241–244, Chengdu, China, August 2003.
[14] ANSI/IEEE Std 802.11, 1999 Edition (R2003), IEEE Standard for Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Network-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999.
[15] IEEE Std 802.11i-2004, IEEE Standard for Information Technology- Telecommunications and information exchange between systems- Local and metropolitan area networks-Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Medium Access Control (MAC) Security Enhancements, 2004.
[16] IEEE Std 802.1X-2001, IEEE Standard for Local and metropolitan area networks Port-Based Network Access Control, 2001.
[17] JPEG 2000 image coding system—Part 1: Core Coding System, ISO/IEC JTC 1/SC 29/WG 1 15444-1.
[18] T. Ebrahimi, C. Christopoulos, and D. T. Lee, Eds., "Special issue on JPEG2000," *Signal Processing: Image Communication*, vol. 17, no. 1, 2002.
[19] T. Ebrahimi and D. D. Giusto, Eds., "Special section on JPEG2000 digital imaging," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 4, pp. 771–888, 2003.
[20] JPEG 2000 image coding system—Part 9: Interactivity tools, APIs and protocols, ITU-T Recommendation T.808, ISO/IEC 15444-9, July 2004.
[21] C. Perra and D. D. Giusto, "A framework for image based authentication," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 2, pp. 521–524, Philadelphia, Pa, USA, March 2005.
[22] K. S. Joo and T. Bose, "Two-dimensional periodically shift variant digital filters," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 1, pp. 97–107, 1996.
[23] Y. S. Sun and H. C. Shyu, "Image scrambling through a fractional GR(q^n) composite domain," *Electronics Letters*, vol. 37, no. 11, pp. 685–696, 2001.
[24] Z. Han, W. X. Feng, L. Z. Hui, L. D. Hai, and L. Y. Chou, "A new image encryption algorithm based on chaos system," in *Proceedings of the IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*, pp. 778–782, Changsha, China, October 2003.
[25] J. Zou, R. K. Ward, and D. Qi, "The generalized fibonacci transformations and application to image scrambling," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 3, pp. 385–388, Montreal, Quebec, Canada, May 2004.
[26] D. Van De Ville, W. Philips, R. Van De Walle, and I. Lemahieu, "Image scrambling without bandwidth expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 6, pp. 892–897, 2004.
[27] S. Lian, J. Sun, and Z. Wang, "A novel image encryption scheme based-on JPEG encoding," in *Proceedings of the 8th International Conference on Information Visualization*, vol. 8, pp. 217–220, London, UK, July 2004.
[28] H. Wu and D. Ma, "Efficient and secure encryption schemes for JPEG2000," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 5, pp. 869–872, Montreal, Quebec, Canada, May 2004.
[29] O. Watanabe, A. Nakazaki, and H. Kiya, "A fast image-scramble method using public-key encryption allowing backward compatibility with JPEG2000," in *Proceedings of the International Conference on Image Processing (ICIP '04)*, vol. 2, pp. 3435–3438, Singapore, Republic of Singapore, October 2004.
[30] JPEG 2000 image coding system—Part 8: JPSEC Final Committee Draft—Version 1.0, ISO/IEC JTC1/SC29/WG1 N 3480, November 2004.
[31] M. Matsumoto and T. Nishimura, "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Transactions on Modeling and Computer Simulation*, vol. 8, no. 1, pp. 3–30, 1998.

G. Ginesu received MS in electronic engineering (2001), discussing a thesis on thermal image processing and pattern recognition, and received his PhD degree in electronic engineering (2004) from the University of Cagliari, Italy. During 2001, he was at the Institute for Telecommunications of the Technical University of Braunschweig, Germany, to work on thermographic image processing. In 2003 he spent a period of 6 months as a Visiting Scholar at Rensselaer Polytechnic Institute, Troy, NY, to work on volumetric data coding (advisory Professor W. A. Pearlman). His research interests are related to image processing and transmission, volumetric data processing and coding, error concealment for wavelet-based image trasmission, and JPEG2000/MPEG standards. He is a Member of IEEE and of the CNIT's Unit of Research in Cagliari.

D. D. Giusto received his MS degree in electronic engineering (1986) and his PhD degree in telecommunications (1990) from the University of Genoa, Italy. Since 1994, he has been a permanent faculty member of the Department of Electrical and Electronic Engineering, University of Cagliari, where he was appointed Full Professor of telecommunications in 2002. He is the recipient of the 1993 AEI Ottavio Bonazzi best paper award, and corecipient of an 1998 IEEE Chester Sall best paper award. Since 1999, he is the Italian Head of Delegation within the ISO-JPEG standardization committee; he is also a Member of the executive board of CNIT, the Italian University Consortium for Telecommunications. He is acting as evaluator/auditor for the

European Commission since 1994. His research interests are in the area of communication systems, multimedia, and video/image processing and transmission. He is a Senior Member of IEEE.

T. Onali received MS degree in electronic engineering in 2004 from the University of Cagliari, Italy, discussing a thesis on image-based authentication. At present, she is pursuing her PhD degree in the CNIT Multimedia Communications Laboratory at the University of Cagliari. Her research interests are related to network and data security, network performance evaluation, and multimedia communications.