ARTICLE IN PRESS

Information Sciences xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Protecting secret documents via a sharing and hiding scheme

Shu-Fen Tu [a,1], Ching-Sheng Hsu [b,*]

[a] Department of Information Management, Chinese Culture University, No. 55, Huagang Rd., Shihlin District, Taipei City 11114, Taiwan
[b] Department of Information Management, Ming Chuan University, No. 5, Deming Rd., Gueishan Township, Taoyuan County 333, Taiwan

ARTICLE INFO

Article history:
Received 25 September 2009
Received in revised form 31 January 2011
Accepted 8 March 2014
Available online xxxx

Keywords:
Document protection
Secret sharing
Steganography
Polynomial interpolation

ABSTRACT

This paper presents a document protection scheme that integrates secret sharing and steganography. The secret document is divided into n shares based on a $k-1$ degree polynomial, and the sharing scheme is tied in with a cover document. With the corresponding cover document, we can recover the secret by gathering at least k shares. To ensure each share is distinct, this work designs a method to avoid duplicate arguments of the polynomial. The security in the proposed scheme comes from two sides, secret sharing and steganography. The cover document is not restricted to the secret, hence our scheme is more flexible in choosing the cover document compared to other researchers' schemes. Besides, the proposed scheme is easy to implement a hierarchical right management. Finally, experimental results and computational and security analysis are given to show that the proposed scheme is secure and practical.

1. Introduction

The Internet facilitates the exchange of information with others. However, data security is a major issue. Several methods can be employed to protect secret information, and these methods adopt various approaches. Classic cryptography protects secret information by scrambling the contents with a private key. Without the appropriate key, unauthorized authors cannot decode the secret within a reasonable amount of time with the available resources. Unlike cryptography, steganography, or information hiding, conceals the secret on an innocent media rather than scrambling it. Another cryptographic method called the secret sharing scheme is proposed to encode the secret into many shares and only by gathering enough shares can somebody recover the secret. Such a scheme can prevent the secret from being leaked out by single person. Due to the advantages and disadvantages of these methods, a few researchers have integrated multiple methods to propose different protection schemes for digital information.

In 1998, Lin and Lee [7] proposed a novel document protection scheme, called Confused Document Encryption Scheme (CDES), which can be seen as an integration of steganography and cryptography. Conveying a secret in plaintext requires picking a clean document, called a cover message, which must include all characters of the plaintext. Take the plaintext "I love you" for example. There are in total eight different ASCII characters — "I", "l", "o", "v", "e", space, "y" and "u", so

* Corresponding author. Tel.: +886 3 350 7001x3741; fax: +886 3 359 3875.
E-mail addresses: dsf3@ulive.pccu.edu.tw (S.-F. Tu), cshsu@mail.mcu.edu.tw (C.-S. Hsu).
[1] Tel.: +886 2 2861 0511x35932; fax: +886 2 861 8805.

http://dx.doi.org/10.1016/j.ins.2014.03.071
0020-0255/

the cover message must include those characters as well. For example, we may take the sentence "I have played guitar for a long time" as the cover message. Each character of the cover message is numbered according to its position, and then each character of the plaintext is encrypted as an integer, which denotes the position index of the same character of the cover message. For instance, the character "l" may be encoded into 9 or 28, which are the indices of "l" in the cover message. Finally, the plaintext becomes a plaintext index file (PIF). The PIF is compressed and encrypted by IDEA [6], and then sent after the cover message is transmitted. However, Lin and Lee's method is only appropriate for languages with small character sets, such as English, and is impractical for double-byte character-based languages and digital images. Yeh and Hwang [11] proposed an enhanced scheme based on CDES to handle messages in any language. They utilized the encoding method for double-byte characters to convert them into hexadecimal inner codes, thus significantly shrinking the character set to the digits '0'–'9' and letters 'A'–'F'. For example, traditional Chinese characters can be converted to hexadecimal inner codes by the encoding method "Big5", and CDES can then be adopted to encrypt these codes. With a little modification, the enhanced CDES can be used for any kinds of digital information, including digital images as well. However, the character set of the cover message must include that of the plaintext; otherwise, the characters absent in the cover message cannot be encrypted. In addition, the cover message should be larger than the plaintext to maintain the diversity of indices in the PIF. The size of the PIF may be very large since the characters are encoded into integers. Moreover, their scheme cannot prevent the single copy of the secret from being betrayed or destroyed.

Some researchers combined the concept of steganography and secret sharing to design protection schemes for digital images [1,2,4,5]. To transform a secret image, their method has to pick n innocent cover images. The main idea behind their methods is to re-draw these cover images according to the secret image, so that each cover image becomes a share of the secret image. To recover the secret image, they have to gather all stego-images and perform the logic OR or XOR operations on these stego-images. Due to the intrinsic nature of their schemes, the stego-images have to be enlarged and appear artificial, and hence may leak out secret information. However, their schemes cannot be applied to documents. In addition, their schemes are mainly a kind of (n,n)-threshold secret sharing scheme, in which a secret is split into n shares and recovering the secret requires all the n shares. Generally, a secret sharing scheme with a flexible threshold for recovering the secret is more desirable. That is, the threshold for recovering the secret can be smaller than the number of shares. Considering the increase in true-color images, Tsai et al. [10] proposed a sharing and hiding scheme for true-color images. Both the secret and cover images are in true color, and the secret is split into n pieces and concealed into n cover images. Their stego-images images do not look artificial, but are enlarged. Their experiments show that the enlarged size is smaller than the results of some other researchers [1,2,8]. However, their scheme is also a (n,n)-threshold secret sharing scheme.

The purpose of this paper is to propose a new document protection scheme. Similar to Lin and Lee's and Yeh and Hwang's schemes, we use a clean document to cover the transmission of the secret document. Unlike their schemes, we will incorporate Shamir's (k,n)-threshold secret sharing scheme [9] to split the secret into many shares, so that we can prevent the single copy of the secret from being betrayed or destroyed. The prominent feature of our method is that the generation of these shares is tied in with the cover document. Therefore, to recover the secret, we have to gather at least k shares and the corresponding cover document jointly. By doing so, we can gain the following advantages. First, a steganography-like way can enhance the security of the scheme. Since the cover document looks innocent, hackers may pass over it and hence lose one of the keys for recovering of the secret. Second, the cover document is not required to be transmitted through the secure channel since it does not contain any secret. It is impossible to acquire information about the secret from the cover document only. Third, integrating secret sharing into our scheme can extend the application of the proposed scheme. For example, we can implement a hierarchical key management scheme, wherein a super user owns a super key, and other users hold a normal key. Accordingly, the cover document can be seen as a super key, and the other shares can be seen as normal keys. Such a hierarchical scheme is very common in financial institutions. Take a legal document, such as a will, as another example. We can split the will into many shares and distribute them to each witness to prevent anyone altering it on purpose. For the sake of generality of the scheme, the inner codes of the document are processed, instead the words directly; therefore, we can handle documents in any language.

The remainder of this paper is organized as follows: Section 2 reviews Shamir's (k,n)-threshold secret sharing scheme proposed in 1979. Section 3 then describes the proposed document protection scheme. Section 4 summarizes the experimental results, and Section 5 presents the computational and security analysis. Finally, some discussions and conclusions are given in Section 6.

2. Shamir's (k,n)-threshold secret sharing scheme

Shamir's scheme splits a secret number into n shares, and requires at least k shares to recover the secret. Splitting a secret number D into n shares $D_1, D_2, \ldots, D_n$ involves randomly picking a prime number p that is greater than both D and n, and constructing the following polynomial of degree $k-1$:

$$q(x) = (a_0 + a_1 x + a_2 x^2 + \cdots + a_{k-1} x^{k-1}) \bmod p, \tag{1}$$

in which $a_0 = D$. Each share of D is evaluated by $D_1 = q(1), \ldots, D_i = q(i), \ldots$, and $D_n = q(n)$. The $k-1$ coefficients $a_1, a_2, \ldots$, and a_{k-1} are randomly chosen from a uniform distribution over the integers in $[0,p)$. Given any k pairs of $\{(i,D_i) \mid i = 1 \ldots n\}$,

Please cite this article in press as: S.-F. Tu, C.-S. Hsu, Protecting secret documents via a sharing and hiding scheme, Inform. Sci. (2014), http://dx.doi.org/10.1016/j.ins.2014.03.071