# Maintaining Information Security in E-Government through Steganology

**Huayin Si**
*University of Warwick, UK*

**Chang-Tsun Li**
*University of Warwick, UK*

## INTRODUCTION

Traditional government structures are sometimes regarded as overly bulky. However, with the rapid expansion of interconnected computer networks and the progressive development of information technology (IT), it is now possible to exchange massive amounts of data at light speed over great distances. These infrastructures and technologies provide the opportunity for governments to transform themselves from huge monsters to compact and efficient organizations. Realizing the potential benefits of IT, as of summer 2004, 198 governments had started their e-government plans to construct digital government based on the Internet (West, 2004).

One of the essential features of e-government is the transmission of confidential information via computer networks. Depending on the sensitivity of the information, the security of some information should be treated at the same level as national security. Although each e-government has its own networks, no government can say no to the Internet, because it would be a waste of resource. However, the Internet is an open environment; therefore, protecting data flowing on the Internet from attacks is a pressing e-government issue.

All governments with such strategies have sought help from cryptographers and devoted huge amounts of both money and time to the development of specially designed information systems and advanced cryptosystems to strengthen information security. Unfortunately, cryptography is not adequate in some applications. As computing power keeps increasing and the techniques of cryptanalysis keep advancing, contemporary cryptosystems cannot and will not work forever. At the 24th Annual International Cryptology Conference (CRYPTO'04), MD5 and a series of related cryptosystems, which are currently in widespread use, were proved unreliable (Wang, Feng, Lai, & Yu, 2004).

From the last decade, *steganology*—the technique for digitally hiding and detecting information – is attracting more attention. It is already regarded as a powerful complement to cryptology and a promising technique for ensuring e-national security. Unlike cryptology, which renders the encrypted information completely meaningless, steganology keeps the host media perceptually unchanged after hiding the secret information. This article will provide an in-depth explanation of the two components of steganology, namely *steganography* and *steganalysis*, and discuss their potential applications in the realm of e-national security.

## STEGANOGRAPHY

Markus Kahn (1995) defined steganography as "the art and science of communicating in a way which hides the existence of the communication. In contrast to cryptography, where the enemy is allowed to detect, intercept and modify messages without being able to violate certain security premises guaranteed by a cryptosystem, the goal of steganography is to hide messages inside other innocent messages in a way that does not allow any enemy to even detect that there is a second message present." This definition has been widely accepted within the information security community.

The application of steganography can be traced back to ancient times. In 499 B.C., Histiaus shaved the head of his most trusted slave and tattooed an important message on his scalp. When the slave's hair grew back, the information was concealed and the slave was sent to Aristagoras, who then shaved the slave's head again to reveal the message, which instructed him to revolt against the Persians (Herodotus, 1992). That probably is the oldest instance of steganography. As the technique improved, invisible ink and microfilm emerged in modern applications. The watermark on bank notes is the most common contemporary instance of steganography.

In the modern world, steganography is a covert communication technology that allows secret information to be hidden in *cover message*/media. The resulting message with the hidden information is called the *stego message*.

Steganographic techniques can be divided into two broad categories: *digital watermarking* and *digital fingerprinting*. Digital watermarking focuses on the embedding algorithms and is used for purposes of copyright protection, authentication and integrity verification. The hidden information, namely, the *watermark*, in digital watermarking is relatively simple, normally the digital signature of the owner or a random pattern generated with a secret key. Digital fingerprinting concentrates on the method (sometimes referred to as protocol) of generating the hidden information, namely, the *fingerprint*, so as to meet requirements such as uniqueness and counterfeit-proofness. Fingerprinting techniques always utilize watermarking techniques to embed the generated fingerprint. In other words, the essential difference between these two categories is that the fingerprint embedded by fingerprinting techniques is unique for every single copy of the cover message, while the watermark used by watermarking techniques is always the same for all copies of the cover message and is related to the cover message and its owner. Different schemes of these two categories also have other special features to meet the specific needs of their applications.

Some other common properties of steganography techniques are as follows:

- **Transparency:** The distortion introduced by the embedding process should be imperceptible to humans so that the impact on the perceptual quality is minimized.
- **Robustness:** For most applications, such as copyright protection, survivability against all kinds of malicious attacks and incidental manipulations, such as lossy compression and format trans-coding, should be maintained unless the manipulations have rendered the content useless in some sense.
- **Payload:** Payload (i.e., the embedding capacity) is important for digital fingerprinting. Since the function of the fingerprint is to identify the individual recipient/buyer, the fingerprint should be long enough to provide space to keep the uniqueness when a huge number of copies of the cover message are to be distributed. In this case, embedding capacity is the deterministic factor of an effective fingerprinting scheme (Su, Eggers & Girod, 2000).

## Digital Watermarking

The idea of digital watermarking is to embed a small amount of secret information—the watermark—into the host media to achieve goals like copyright assertion, authentication and content integrity verification, and so forth. The superiority of digital watermarking over cryptography is that the latter provides no protection after the content is decrypted, while the former provides "intimate" protection at all times, because the watermark has become an inseparable constituent part of the host media. All the capabilities of a watermarking scheme, including the balance between the transparency and robustness to avoid any perceptible artefacts and the other properties to meet its special application, are dependent on the design of the embedding algorithm. To optimize performance, the embedding algorithm is always specially designed for a certain type of media, such as image, video, audio and so forth, to avoid any possible security gap. Digital watermarking schemes can be classified into three categories: *robust watermarking*, *semi-fragile watermarking* and *fragile watermarking*.

Robust watermarking is intended for the applications of copyright protection and digital rights management (DRM), wherein the watermark containing copyright information should be detectable after attacks that aim at erasing the watermark but maintaining the value of the host media. Cox, Kilian, Leighton, and Shamoon (1997) proposed the concept of spread-spectrum (SS) watermarking, which has inspired a great number of recent works in this field. Adopted from communication theory, the idea of SS watermarking is to treat the low-energy watermark as a narrow-band signal and spread it into multiple components in the spectrum of the host media, which is treated as a wide-band signal. By spreading the watermark in the spectrum, the energy of the watermark in a signal frequency is limited, and thus the robustness is guaranteed even when some frequency components are missing. However, the high robustness of SS watermarking is gained at the expense of low payload, so it is not quite suitable to the purpose of digital fingerprinting. To further improve the robustness without causing more artefacts, human perceptual models (HPM), including human visual system (HVS) and human auditory system (HAS), have been proposed and incorporated in the watermark embedding process (Barni, Bartolini & Piva, 2001). Feasible perceptual models facilitate adaptive watermark embedding in components where HPM is less sensitive.

Semi-fragile and fragile watermarking have been developed for the purposes of authentication and content integrity verification, in which the embedded watermark is expected to be destroyed when the attacks are mounted, so that the alarm will be raised by the detector when it fails to extract the watermark. The difference between these two sub-categories is that semi-fragile watermarking regards some designer-specified operations as non-malicious actions while fragile watermarking treats all kinds of manipulations as malicious. Counterfeit-proofness is a key objective (semi-) fragile watermarking schemes are expected to achieve. Counterfeiting attacks, such as cut-