Secure Information Hiding in SOAP Messages Using Cryptography and Steganography

Lwin Mar Thin, Nan Sai Moon Kham
University of Computer Studies, Yangon
lwinmarthin85@gmail.com,moonkhamucsy@gmail.com

Abstract

Computer techniques are progressing remarkably and the amount of electronically transmitted information is increasing every day. Different security measures can be employed to ensure the integrity and confidentiality of secret information. As a result, steganography has become an interesting and challenging field of research striving to achieve greater immunity of hidden data on the host cover media like image, audio, or text. Although techniques for images or sounds have mainly been studied, there are few examples of research of the information hiding method on text data. The proposed system uses an SOAP message as cover text, embedded secret information using stego-key. Steganography especially combined with cryptography is a powerful tool which enables people to communicate without possible eavesdroppers even knowing there is a form of communication in the first place. It provides double security. Digital Signature which is a cryptographic technique is used to authenticate messages between parties. In this study, we propose a secure and undetectable steganography technique to be used for SOAP signature messages within XML Web services environments. This technique is embedded secret data into an innocuous cover text by shuffling XML elements to form the stego text with minimum degradation.

Keywords - SOAP, information hiding, steganography, XML, digital signature

1. Introduction

With the rapid development of Internet technologies, the amount of information sent and received electronically is increasing greatly. The need for private and sufficiently secure communications in several applications such as e-banking, e-trading, mobile telephony, medical data interchanging etc., is rapidly increasing[6]. Different security measures can be employed to ensure the integrity and confidentiality of secret information. By combining steganography and cryptographic technique which enables people to communicate without attack by eavesdroppers. It provides double security. An intruder can break the key for cryptographic or can find out Steganographic techniques but breaking the combination of both can be nearly possible [1].

Information hiding is a field of information security, and it can be applied in XML documents which constitutes towards the field of steganography. Figure 1 shows the general model of the information hiding. To develop the methods of information hiding using XML makes realize the way to establish secret communication channel using SOAP messages, and the ability to trace the source of unauthorized copied [14]. Although, compared with the information hiding methods intended for images and sounds, there are few methods of hiding information into text. Unfortunately, there is almost no study on the methods for hiding information in structured documents.

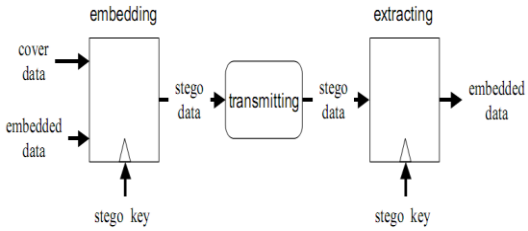


Figure1. Information hiding

Text steganography refers to the process of hiding secret information in text files. For security and imperceptibility reasons, it is very important for stego texts not to show any detectable artifacts. Thus, readers should not notice or discover the modifications made in the stego text files. Generally, the redundant information in text files is very limited in comparison to that in images and audio files. Therefore, using text as cover files in steganography represents the most difficult way of information hiding [13]. Unlike the steganography methods used with conventional covers such as digital images, there are very few steganography methods that can feasibly be used with communication protocols such as SOAP. Thus, this paper also aims to find out an undetectable steganography method based on SOAP messages.

The rest of the paper is organized as follows: section 2 reviews the related work on text and XML steganography. Section 3 describes about SOAP message. Section 4 discusses and explains the security concept of steganography and XML digital signature. Section 5 presents information hiding in SOAP Message. Section 6 describes the proposed technique and gives detail on shuffling of XML tags. Section 7 is discussed in analysis of the results. Section 8 concludes the paper.

2. Related Work

Most research of steganography was using cover media such as images, videos and sound. However, steganography into the text is usually not preferred because of the difficulty in finding redundant bits in a text document [12]. Some of the methods proposed to solve the problem, such as by line shifting, words shifting, up to whitespaces manipulation into the cover text [4].

Por and Delina [7] improved the techniques for data hiding proposed by [14]. Therefore, they proposed a hybrid steganography method for text by combining both inter-word spacing and inter-paragraph spacing methods. Thus, whitespaces between words and paragraphs in right-justification of text are used for data hiding in order to increase the embedding capacity. However, the cover text was dynamically generated according to the size of the secret message.

Banerjee et al (2011) conducted research on text steganography, on his paper [2]. Banerjee introduces a new method of hiding messages into the text by changing the prefix “a” or “an” into the cover of the English text. Shirali-Shahreza [10] proposed a new steganography method for texts. This method is based on the different spelling of some words in English between UK and US. For example, “centre” has different terms in UK (centre) and US (center).

The model proposed in [8] defines a text steganography method based on substituting the words which have different terms in UK and US. For example, (Gas) has different terms in UK (Petrol) and US (Gas).

Liu et al. [9] proposed a text steganography method to be used in online chat. This method is based on an Internet meme named typoglecymia, which means that changing the order of word’s middle letters has a slight to no effect on the ability of skilled readers to understand the text (e.g. Guitar and Guiatr). Therefore, it used the redundancy found in the interior letters’ order. Since this letter randomization equals to the common error made by chatters due to high speed typewriting, it is likely to be used in online chats, where the text usually contains mistakes.

However, the previous studies provided text steganography method, which are not necessarily applicable in SOAP messages context due to the fact that SOAP messages are exchanged and monitored by computer systems rather than humans. Using misspelled or alternative words in SOAP messages would result in the SOAP parsers not being able to handle the SOAP messages received because they do not comply with the expected semantic. In 2011, Although Almohammad ,Alrough, and Ghinea[3] proposed a steganography method to be used for SOAP messages within web services, only steganography method had not fully secured to this environment. The attacker who had known stego key will easily eavesdropped the secret message. This paper proposed a shuffling XML Tags technique. In this process, XML tags are Shuffling can be done multiple times in order to achieve more security. Moreover, a

cryptographic technique is used to authenticate messages between parties .Furthermore, the stego SOAP message has the same size of the original message. This technique is simulated using a feasible scenario so as to demonstrate its utility and applicability.

3. SOAP Message

Communications and interactions between clients and the web service are achieved using Simple Object Access Protocol (SOAP) which supports a common data transfer protocol for effective communication over the Web. Thus, structured and typed information can be exchanged between peers of distributed environment using SOAP messages. A SOAP message is an XML document created in a specific format and it mainly consists of envelope, header, body and fault elements, as shown in Figure 2.

```
<?xml version="1.0" encoding="UTF-8"?>
<S:Envelope
xmlns:S="http://schemas.xmlsoap.org/soap/envelope/">
  <S:Header>
       …
  </S:Header>
  <S:Body>
       …
 <S:Fault>   …
</S: Fault >
  </S:Body>
</S:Envelope>
```

Figure2. SOAP Message Structure

In Web Services, the interaction between service providers and requesters occurs via XML-based SOAP messages. Therefore, such messages offer a kind of steganography cover files. Hence, secret information can be embedded in SOAP messages and sent over the network to an intended destination. Essentially, a SOAP message is a kind of XML document, yet XML documents essentially contain text since they represent a formatted form of text files. Therefore, steganography methods used for text files and XML documents can theoretically be used for SOAP messages. Practically, some or all of these methods might be infeasible since the SOAP protocol has its own structure and application.

In conclusion, SOAP messages are extensively used over the Web and so represent a valuable kind of steganography cover. Thus, it would be a good idea and practice to find out a steganography method that can be used for SOAP messages which would not attract the attention of attackers.

4. Security techniques

To make the SOAP messages secure and less exposed to vulnerabilities, Steganographic technique comes in picture. This procedure reduces the risk of security holes of unknown/undocumented fields that might otherwise compromise the resources. Using Steganographic techniques, the sender and receiver do not have to worry about any secret key and transmitting it from sender to receiver. At the same time, these techniques consume less time and overhead resources. These Steganographic techniques for SOAP can be applied along with cryptographic techniques to get more security. An intruder can break the key for cryptographic or can find out Steganographic techniques but breaking the combination of both can be nearly impossible.

4.1 Steganography

Internet users frequently need to store, send, or receive personal and confidential information. Therefore, such data needed to be kept secure between or during traversing through the network. One of the many methods used to keep data secure is known as steganography. Digital steganography is the art of inconspicuously hiding data within data. Steganography, coming from the Greek words “stegos”, meaning “roof or covered” and “graphia” which means “writing”, is the art and science of hiding the fact that communication is taking place. Using steganography, a secret message can be embedded inside a piece of unsuspicious

information and can be sent without anyone knowing of the existence of the secret message [11].

Steganography is a means of storing information in a way that hides that information's existence. Paired with existing communication methods, steganography can be used to carry out hidden exchanges. Secrets can be hidden inside all sorts of cover source: text, images, audio, video or multimedia [5].

Figure3. Types of Steganography

Cover source is the carrier of message. Most teganographic utilities nowadays, hide information inside images, as this is relatively easy to implement. However, there are tools available to store secrets inside almost any type of cover source. It is also possible to hide information inside texts, sounds and video films [10]. The most important property of a cover source is the amount of data that can be stored inside it, without changing the noticeable properties of the cover. The advantage to prefer text steganography over other media is its smaller memory occupation and simpler communication.

4.2 XML Digital Signature

XML Signature is an electronic signature technology which is defined to be used in XML data transmission. XML signature specification defines electronic signature formats using XML, the creation of electronic signature and rules for the verification process. It solves security problems such as authentication¸ integrity and non-repudiation. Further to this XML Signature provides advanced benefits such as

- Partial signature; allows only data contained in specific tags to be signed in the XML document
- Multiple signature; enables multiple electronic signature to be included in the XML document

XML digital signature used for the verification of claims by the sender should be confirmed by requiring the sender to prove knowledge of the corresponding private key. This can be established through a challenge response mechanism whereby the relying party sends a message that the sender is required to sign and subsequently send back for verification. Once again, timestamps, sequence numbers or expirations should be used to ensure the integrity of these messages and to prevent replay attacks. The example format of digital signature is shown in below.

```
<Signature>
   <SignedInfo>
      <CanonicalizationMethod/>
      <SignatureMethod/>
      (<Reference URI? >
         (<Transforms>)?
         <DigestMethod>
         <DigestValue>
      </Reference>)+
   </SignedInfo>
   <SignatureValue>
   (<KeyInfo>)?
   (<Object Id?>)*
</Signature>
```

5. Information Hiding in SOAP Message

Hiding secret information in a SOAP message means that the mule that is used to convey the secret message is the communication protocol that governs the actual path over a network, instead of using the actual data itself as a cover. This idea can overcome many of the limitations that faced the conventional steganography techniques. Traditional techniques hide secret messages inside digital files, which impose the threat of detecting the secret as these files are usually saved. Alternatively, a SOAP message leaves almost no trail as they are

normally deleted after receiving the message. In addition, a secret piece of information can be divided into multiple smaller messages and transmitted over several SOAP messages to overcome the size limitation as well. Figure 4 illustrates the general model of data hiding in SOAP messages.

When the stego SOAP message arrives at the receiver endpoint, the secret message is extracted using a stego key that is shared between the sender and receiver.

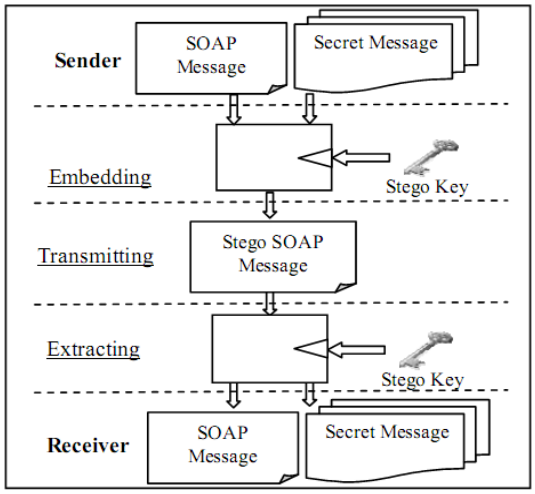


Figure4. SOAP Steganography Model

5.1The Proposed Steganographic Technique using Shuffling XML Tags

Encryption can be used to preserve data security but the technologies required for encryption cause problems with firewalls and they don't work very well on the Internet. Encryption has another problem; if both communication parties don't have the same platform then the receiver can't decrypt the sender's message. As a result, SOAP based steganography method could be a reasonable solution for transmitted data security. Steganographic techniques have been designed which ensure varying degree of security for XML documents. The provided method has a high resistance against detection since it uses the communication protocol as a cover medium rather than the traditional digital files. Figure 5 shows the process flow of Shuffling of Tags Technique.

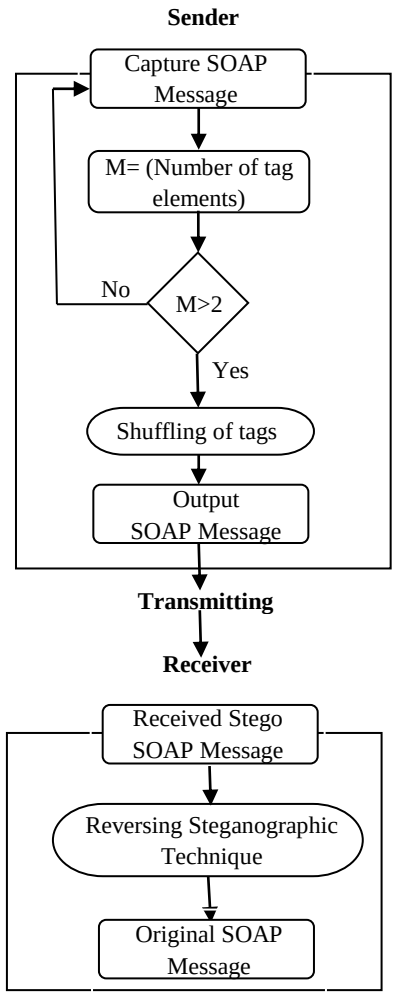


Figure5. Working of Shuffling XML Tags

The embedding procedures of shuffling XML tags are as follows:

Step1.Capture the cover SOAP message to embed secret data.

Step2.Scan the whole XML file tag by tag.

Step3.Calculating the number of tag elements (M) of the document that can be used to hide data (N).

Step4.If all the symbols of the secret message can be hidden in one SOAP message (the number of available sets of N is smaller than the tags of the secret message (M), and then continue the following steps. Otherwise, another SOAP message should capture and repeat again step 2.

Step5.Swap the 1st tag with the last (nth) tag of the document.

Step6.Swap the 2nd tag with second last tag (nth -1).

Step7.Recursively apply the procedure of swapping tags throughout document.

The extracting procedures of the receiver are as follows:

Step1. Select the Stego SOAP message.

Step2.To obtain original document, at the receiver end, this process is reversed and shuffling of tags again taken place to restore the original sequence of tags.

In this process, XML tags are shuffled. This shuffling is done in a sequence. The position of 1st tag along with its value, are exchanged with the last tag and its value. The second tag is shuffled with the second last tag of XML file. This process repeats itself until; it reaches the end of file.

6. Discussion

There are mainly three aspects that should be taken into account when discussing the results of the proposed method of information hiding. They are security, capacity and robustness. The original message and stego message results are shown in the figures 6 and 8. This method satisfies security aspects and robustness. It generates the stego message with minimum degradation which is not very revealing to people about the existence of any hidden data, maintain its security to the eavesdropper. Besides the security level has increased through the signing of the XML document after embedding operation. But the capacity of this method depends on the capture SOAP message. This technique works best for large XML files. In this system, we used CD catalog order file as an innocuous cover text for example scenario and embedded data to send different sequences of order to the receiver.

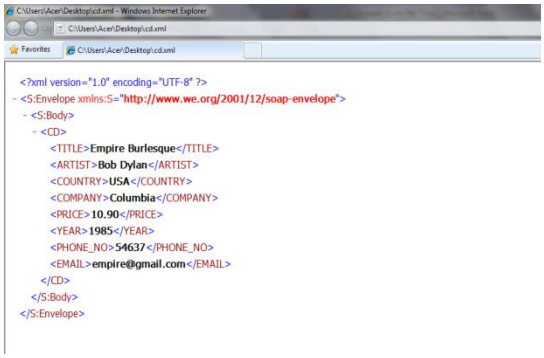
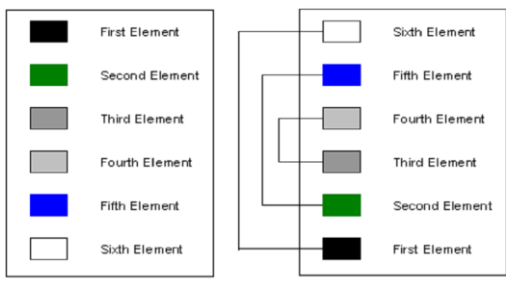


Figure6. Cover SOAP message

In figure 7, XML file is represented in Part A. The six colored boxes show XML elements. These elements can be referred as tags. Figure 5, Part-B shows that these XML elements are shuffled with each other. These figures depict the shuffling of XML tags.



Part A Part B

Figure7. Model for shuffling of Tags

Secret Message: “Hey”

Stego Key: “H”= shuffle(1st, 8th)element
“e”= shuffle(2nd, 7th) element
“y”=shuffle(3rd,6th)element

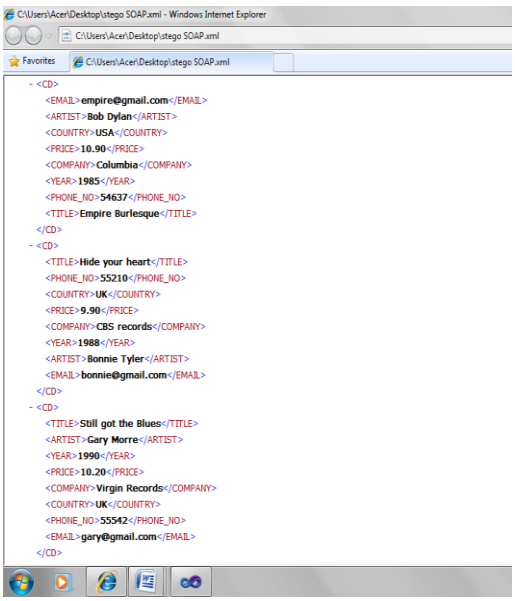


Figure8. Stego SOAP message

7. Conclusion

In the study of Distributed XML Web Services using Steganography Techniques, different steganography techniques are applied on different samples of XML documents. Moreover, digital signatures are added to this system. So this ensures benefit of both the fields and thus can present an XML document with more security, to exchange curial information in this requisite era of Internet technologies. The proposed techniques, as demonstrated by examples, clearly hide information in SOAP message and, in such a manner that concerned receiver alone can access the information. The technique of shuffling of tags is best for large file and tags can be shuffled multiple times to achieve more security. This technique does not increase file size and provides security.

References

[1]Alrouh, B., Almohammad, A., and Ghinea,G., “Information Hiding and SOAP Messages”, The 5th International Conference (ICITST 2010), London, UK, 8-11 November, 2010.

[2]Banerjee, Indradip, S.Bhattacharyya and G. Sanyal, “Novel Text Steganography through Special Code Generation”, International Conference on Systemics, Cybernetics and Informatics, Pentagram Research Centre (P) Limited,2011.

[3]B.Alrough, A.Almohammad and G.Ghinea, “Information Hiding in SOAP Messages: A Steganographic Method for Web Services”, International journal for Information Security Research(IJISR), vol.1, Issue 1,2011

[4]Carro, Fernando Incertis (2007) “Methods of invisibly embedding and hiding data into soft-copy text documents”, U.S. Patent No. 7240209 B2 July 3rd ,2007.

[5]G.Davida, M.Chapman and M.Rennhard, “Effective approach to large-scale automated linguistic steganography”, In Proceedings of the Information Security Conference, pages156–165, October, 2001.

[6]Klimis S.Ntalianis, “A Short-Message Robust Steganographic Method for Effective Information Recovery Under Transmission Losses of Cellular Networks”, Proceedings of the 9th WSEAS International Conference on Systems, Greece, 2005, pp. 955-957.

[7]L. Y. Por and B. Delina, “Information Hiding: A New Approach in Text Steganography”, In Proceedings of the 7th International Conference on Applied Computer & Applied Computational Science (ACACOS’08), Hangzhou, China, 2008

[8]M. H. Shirali-Shahreza and M. Shirali-Shahreza, “A New Synonym Text Steganography”, In Proceedings of the 4th International Conference on Intelligent Information Hiding and Multimedia

Signal Processing (IIH-MSP 2008), Harbin, China, 2008.

[9]M. Liu, Y. Guo, and L. Zhou, “Text Steganography Based on Online Chat”, In Proceedings of the 5th International Conference on Intelligent InformationHiding and Multimedia Signal Processing (IIH-MSP 2009), 12-14 Sep, 2009, Kyoto, Japan, 2009.

[10]M. Shirali-Shahreza, “Text Steganography by Changing Words Spelling”, In Proceedings of the 10th International Conference on Advanced Communication Technology (ICACT 2008), Phoenix Park, Korea, 2008.

[11]R. Krenn. Steganography and Steganalysis, Proceedings of IEEE Conference,University of California, pp 142, 2004

[12]Singh, Hitesh, Pradeep Kumar Singh and Kriti Saroha (2009) “A Survey on Text Based Steganography”, Proceedings of the 3rd National Conference, Computing For Nation Development, February 26 – 27, 2009.

[13]S. Inoue, K. Makino, I. Murase, O. Takizawa, T. Matsumoto, and H. Nakagawa, “A Proposal on Information Hiding Methods using XML”, In Proceedings of the 1st Workshop of NLP and XML, Nov, 2001.

[14]W. Bender, D. Gruhl, N. Morimoto, and A. Lu, “Techniques for Data Hiding”, IBM Systems Journal, vol. 35(3-4), pp. 313-336, 1996.