

STEGANOGRAPHY IN WIRELESS APPLICATION PROTOCOL

Mohammad Shirali-Shahreza
Computer Science Department
Sharif University of Technology
Azadi Street
Tehran
Iran
shirali@cs.sharif.edu

ABSTRACT

Mobile phone and Internet technologies have progressed along each other. The importance of both these technologies has resulted in the creation of a new technology for establishing wireless Internet connection through mobile phone, known as Wireless Application Protocol (WAP). However, considering the importance of the issue of data security and especially establishing hidden communications, many methods have been presented. In the meanwhile, steganography is a relatively new method.

In this paper, a method for hidden exchange of data has been presented by using steganography on WML pages (WML stands for Wireless Markup Language, which is a language for creating web pages for the WAP). The main idea in this method is hiding encoded data in the ID attribute of WML document tags. The coder program in this method has been implemented using the Java language. The decoder program to be implemented on the mobile phone has been written with a version of Java language specifically used for small devices, which is called J2ME (Java 2 Micro Edition). It was tested on a Nokia series 60 mobile phone.

KEY WORDS

Information Hiding, Internet Security, WML, WAP, Steganography, and Mobile Phone

1. Introduction

With the advent of the Internet and its rapid progress, all aspects of daily human life were unbelievably influenced by the Internet and the human strongly depended on the Internet. Internet importance resulted in considering using the Internet in all places, even outside home and workplace, so the wireless Internet has been created.

Along with the expanse of the Internet, mobile telephones are increasingly welcomed and mobile phone manufacturing companies have been trying to improve and expand the use of mobile phones.

The synchronous appearance of these two events has created the idea of the wireless Internet for the mobile phone. After many years of disorder application of

technology in establishing wireless communication, finally the WAP Forum was formed in December 1997 by Nokia, Ericsson, Motorola and Unwired Planet companies for the cause of creating a single standard for wireless communication and distant services on mobile phones and other wireless devices [1]. The WAP 2.0 was released in January 2002 [2]

The WAP (Wireless Application Protocol) introduces a communication protocol and the programming environment for implementing web-based information systems within the framework of devices such as pagers, pocket PC's and mobile phones. WAP has introduced a markup language for description of information, which is known as the WML (Wireless Markup Language). It is a XML-based language and, in fact, replaces the HTML formatting language and other page formatting languages such as XML [3]. The reasons for this can be summarized as follows:

In desktop computers, web pages are read, processed and finally displayed by browsers. However, mobile phones, because of the small size of the screen, low bandwidth and a weak processor, are not able to display web pages [4].

Also for other reasons such as limited input devices (lack of a mouse or keyboard), limited memory and limited battery power, it is not possible to display complicated web pages such as HTML pages on mobile phones. Therefore, it becomes necessary to create a substitute language for web pages in mobile phones. By using WML, one can provide data to user optimally [5, 6, 7, 8].

Besides all the points mentioned so far, the issue of data security has become increasingly important, especially in establishing wireless communications, in which there is the possibility of disclosure of confidential and/or personal information during exchange of data between different systems. One of the important branches in the field of information security is the issue of hidden data exchange. For this purpose, various methods such as cryptography, steganography and coding, were used.

Steganography is a method considered in recent years. The main purpose of steganography is hiding information in other cover media, so that other persons will not notice hidden information. This is the main distinguishing attribute of this method with the other methods of hidden data exchange, because, for example in the cryptography

method, individuals seeing the coded data will notice the data but they will not comprehend them. However, in the steganography method, the individuals will not notice data in the sources at all [9].

Most steganography jobs have been carried out on images [10, 11], video clips [12, 13], music and sounds [14]. However, very little work has been done on text data steganography [15, 16].

Nowadays, information security has been considerably improved by combining the steganography method with the other methods. In addition to application in hidden exchange of information, steganography is also used in other fields such as copyright, preventing e-document forging, etc [15, 16].

Considering the above points, this article presents a new method for steganography of information on WML pages.

2. Previous Works

The WML is based on the XML (Extendible Markup Language). The XML was made based on the HTML [17]. As no activity on WML file steganography has been reported, we deal with activities carried out on information steganography in HTML files, which are very similar to WML files.

2.1 Hiding Data in Comments [18]

One can add comments on HTML pages as with other languages. On HTML pages, the comments begin with `<!--` and end with `-->`. An easy way of steganography on HTML pages is to place data as comments.

As the browser does not show this information, the regular user cannot view these comments. However, in the WAP technology, before sending the WML file from the server to the user, first the information is sent to a gateway known as the WAP Gateway, which compiles the WML file in binary form. In this compiled page, all additional page comments and white spaces will be removed and, in general, any unnecessary code is left out of the page. Now the compiled page is sent to the mobile phone through the WAP. The micro-browser (which is a browser used for the mobile phones) processes and displays the received page [3].

Therefore, considering all these points, one cannot employ steganography in WML comments.

2.2 Steganography by Creating White Spaces

On HTML pages, the existence of spaces in a tag is not important and the browser does not pay attention to these spaces. This feature can be used for hiding information by placing certain white spaces among tag members. Then, while extracting information from the page, information is extracted by calculating the amounts of spaces and by using the appropriate function.

As mentioned in section A, while compiling WML files, all additional spaces are removed. Therefore, this method is not applied here either.

2.3 Hiding Data by Changing the Case of the Letters

Another feature of HTML documents is their case-insensitivity of tags and their members. For example, the three tags `
`, `
` and `
` are equally valid and are the same. As a result, one can do information steganography in HTML documents by changing the small or large case of letters in document tags. To extract information, one can extract information by comparing these words with words in normal case and by using the appropriate function.

However, in the WML, all tags should be written in lowercase letters and, as a result, this method cannot be employed.

2.4 Hiding Data at the End of File

HTML pages begin with `<html>` and end with the `</html>` tag. One of the other simplest methods for information steganography on HTML pages is to place them after the `</html>` tag at the end of the page. As the browser does not show such information, the user cannot see the information.

As it was mentioned in section A, while compiling WML files, the additional codes are removed. Therefore, all texts after `</wml>` (which is equivalent to the `</html>` tag on HTML pages) are removed. As a result, this method cannot be applied either.

2.5 Steganography by Reordering Attributes of Tags [19, 20]

One of the features of HTML documents is the insensitivity of these documents to the order of the tag members. Therefore, by using this, one can hide information in HTML documents.

To extract information from the document, the order of the tag members is compared with the sorted state of the tag members and the hidden information is thus extracted. After compiling the WML pages at the gateway, it is not possible anymore to identify the order of the tag members. Therefore, this method cannot be implemented either.

3. Suggested Algorithm

One of the attributes that exist almost in all tags on WML pages is the ID attribute. Based on this attribute, each tag is given a unique ID code. Our purpose in this algorithm is to hide information in this tag attribute. To this end, first the message in question is converted into a coded phrase by the following function:

3.1 Coding Function of the Main Phrase:

This function changes each byte into two characters. Each byte has a code from 0 to 255. To convert this byte to two characters, first we save the unitary digit. The two tens and hundreds figures are in total between 0 and 25, which corresponds exactly with the number of English letters A to Z. Therefore, we save the corresponding letter of these two digits. We repeat exactly this action for the subsequently bytes so as to code the entire phrase. (Fig. 1)

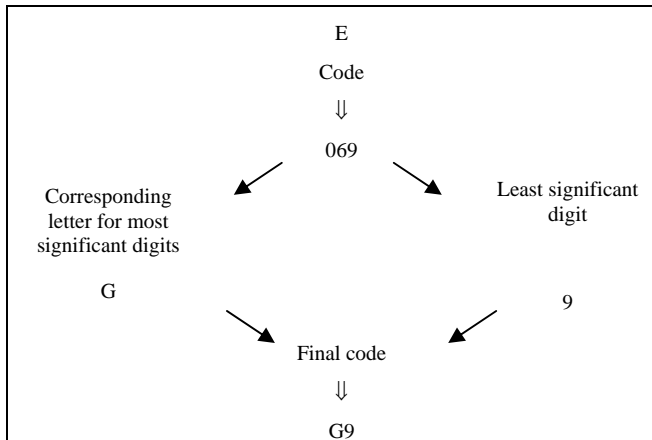


Fig. 1

An example of the function for conversion of a byte into two characters

Now, we choose the tags which have no ID attribute. Then an ID code with the following amount is allocated to them. First, the amount of "i_" is put as a sign and then a zero (the reason is mentioned later). In the end, we put two characters of the coded string. The obtained amount (which is a five characters string) is put as the ID code amount.

The reason for putting zero after two sign letters is to prevent the similarity of the two ID codes. If the two ID codes were exactly similar, in the second name, instead of the digit zero, a digit one is put. In the same way, if more ID codes were similar, the digits will be incremented.

Care must be taken to place the ID attributes only on the tags that do not contain ID attribute. Therefore, the amount of information able to be hidden is variable and does not depend just on the number of the tags. It also depends on the number of tags without ID attribute to which ID attribute can be added.

When decoding, as the pages received on the mobile phone have been compiled, it is not possible to read the WML page tag by tag or to extract the coded character. However, the considerable point is that, on compiled pages, the amount allocated to the ID attribute has been saved in the file without changing. Therefore, to find the amounts hidden on pages, first we search the sign put in the beginning of the ID code as sign, i.e. "i_" and, by finding this sign, save the two successive characters after this sign as the coded characters. In the end, by putting the coded characters besides each other, the coded phrase is obtained. Then, by using the following function, the original phrase is obtained.

3.2 Decoded Function of the Coded Phrase:

First, the corresponding digit is identified with the English letter. (Letters A to Z correspond to digits 0 to 25, respectively). This digit is put behind the previous digit. The obtained number (which can be a 1- to 3-digit number) is the code of the byte in question. The same action is carried out on other letters and digits so as to extract information (Fig. 2).

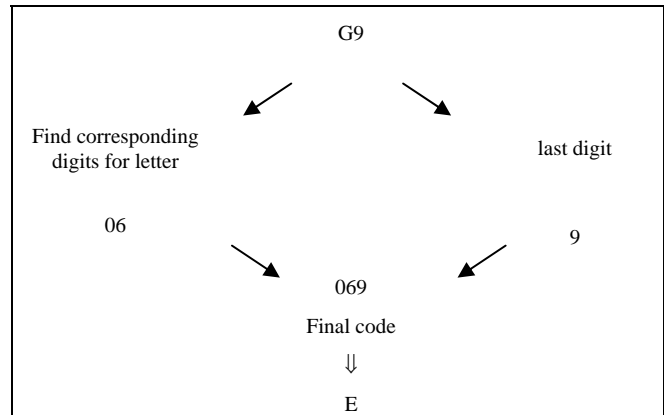


Fig. 2

An example of the function for converting two characters into a byte

4. Advantages and Disadvantages

4.1 Advantages

- Because of the design of the mobile phone systems, access to saved pages is very difficult. Therefore, one cannot easily find the saved WML pages and, by seeing inside the file, he can not access the hidden information.
- The ID attribute is one of the commonly used attributes in tags. Therefore, even in the event of seeing the source of WML file, it will not attract attention.
- Because of compiling of the pages at the gateway, it is very difficult to extract information from the page.

4.1 Disadvantages

- Use of the Internet in mobile phones is not yet used widely by the public [4].
- Because of the difference in the mobile phone Operating Systems, the decoding program may not be properly run on other mobile phones.
- Because of the small size of the WML pages, the volume of information hidden on these pages is very low.
- In this method, the size of WML file is increased and the page transmission speed is reduced.

5. Experimental Results

In this project, by using the described algorithm, the steganography of the messages and phrases on WML pages were dealt with. To this end, some messages were selected. After that, several WML pages were selected and the capacity (volume of information that can be hidden in it by steganography) of each page was calculated and the pages that could undergo message steganography were selected. Then, with the help of the steganography program-which is written in Java and is run on the server-the messages were hidden on the WML pages. ID attributes of the WML files were used for hiding data in proportion to the described algorithm and had the intended format.

After that, with the help of the decoded program, which was written in J2ME (Java 2 Micro Edition, which is a language for small devices such as the PDA's and mobile phones) and was run on the mobile phone, steganography messages are extracted from WML files. By comparing the extracted messages with the input messages, it was seen that both messages were identical and the results were satisfactory. We tested the decoder program on a Nokia series 60 mobile phone.

6. Conclusion

In this article, a new method for information steganography in web pages based on WAP technology (on WML pages) was dealt with. With this method, one can exchange hidden information through the Internet and with the help of the mobile phone.

By using a scripting language on the server side, such as ASP.NET or PHP, one can expand the used method to produce dynamic web pages and hide data in these dynamic pages [21].

Since WAP can be used in other wireless devices such as PDA and pagers, so we can apply this method for them. Considering the increase in wireless Internet speed and the development of mobile phone capabilities throughout the world, more improved methods for mobile phone steganography can be applied in the future.

References

- [1] D.D. Paro, Wireless Application Protocol (WAP): What is it all about....How does it work, *SANS InfoSec Reading Room Wireless Access*, September 4, 2001.
<http://www.sans.org/rr/whitepapers/wireless/148.php>
- [2] WAP Forum, All specifications belonging to the WAP 2.0 release,
http://www.openmobilealliance.org/tech/affiliates/wap/technical_wap2_0.20021106.zip
- [3] WAP Forum, Wireless Application Protocol 2 – Technical White Paper, January 2002
<http://www.wapforum.org>
- [4] D. Viehland, and J. Hughes, The future of the wireless application protocol, *Eighth Americas Conference on Information Systems, AMCIS 2002*, 2002, 1883-1891.
- [5] M. METTER, and R.M. COLOMB, WAP Enabling Existing HTML Applications, *Proc. of First Australasian User Interface Conference Canberra, AUIC 2000*, Australia, 31 January – 2 February 2000, 49-57.
- [6] E. Kaasinen, M. Aaltonen, J. Kolari, S. Melakoski, and T. Laakko, Two approaches to bringing Internet services to WAP devices, *Proceedings of the 9th international World Wide Web conference on Computer networks: the international journal of computer and telecommunications networking*, Amsterdam, The Netherlands, 2000, 231 - 246.
- [7] L. Choi, D. Kim, S. Lee, and C. Kang, Adaptive and Automatic Creation of Hierarchical WML Decks for Efficient Access of Wireless Internet under Wireless Application Protocol, *CDMA International Conference*, 2002, 453-460.
- [8] L. Chittaro, and P. Dal Cin, Evaluating Interface Design Choices on WAP Phones: Navigation and Selection, *Personal and Ubiquitous Computing*, 6(4), 2002, 237-244.
- [9] J.C. Judge, Steganography: Past, Present, Future, *SANS white paper*, November 30, 2001.
<http://www.sans.org/rr/papers/index.php?id=552>
- [10] R. Chandramouli, and N. Memon, Analysis of LSB based image steganography techniques, *Proceedings of the International Conference on Image Processing*, vol. 3, 7-10 Oct. 2001, 1019 - 1022.
- [11] M. Shirali Shahreza, An Improved Method for Steganography on Mobile Phone, *WSEAS Transactions on Systems*, 4(7), July 2005, 955-957.
- [12] G. Doërr, and J.L. Dugelay, A Guide Tour of Video Watermarking, *In Signal Processing: Image Communication*, 18(4), 2003, 263-282.
- [13] G. Doërr, and J.L. Dugelay, Security Pitfalls of Frame-by-Frame Approaches to Video Watermarking, *In IEEE Transactions on Signal Processing, Supplement on Secure Media*, 52(10), 2004, 2955-2964.

- [14] K. Gopalan, Audio steganography using bit modification, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '03)*, vol. 2, 6-10 April 2003, 421-424.
- [15] N. F. Maxemchuk, and S. Low, Marking Text Documents, *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, USA, Oct. 26-29, 1997, 13-16.
- [16] A.M. Alattar, and O.M. Alattar, Watermarking electronic text documents containing justified paragraphs and irregular line spacing, *Proceedings of SPIE -- Volume 5306, Security, Steganography, and Watermarking of Multimedia Contents VI*, June 2004, 685-695.
- [17] W3C, Extensible Markup Language (XML), 2005/08/02.
<http://www.w3.org/XML/>
- [18] HIPS Systems, ShadowText, last visited: 29 July 2005.
<http://home.apu.edu/%7Ejcox/projects/HtmlStego1>
- [19] S. Forrest, Introduction to Deogol, last visited: 26 July 2005.
<http://wandership.ca/projects/deogol/>
- [20] J. Corinna, Steganography 13 - Hiding binary data in HTML documents, last visited: 26 July 2005.
<http://www.codeproject.com/csharp/steganodotnet13/>
- [21] B. Webber, Using Microsoft's ASP.NET Mobile Controls to Develop a Truly Usable Mobile Web Application, Dissertation, Computer Science Department, Rhodes University, November 2003.