Volume 3, Issue 11, November 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Use of Stegnography in Hiding Text Using CSS in Markup Language

Kapil Kumar Kaswan*
Research Scholar
Department of Computer Science & Studies
Mewar university,Chittorgarh,Rajasthan,India

Dr. Roshan Lal
Assistant Professor
Department of Computer Science
Govt College,Karnal, India

Abstract-Text stegnography is used to hide the message behind some other cover text.we will use the html tags,css and their attribute to hide the secret message .it is based on the fact that the ordering of the attribute in the html tags has no impact on the appearance of the document. The css can contain some selector that are not used regular can hide the information.this ordering can be used to hide the secret message efficiently.the proposed technique has two component that are hiding process and extracting process.hiding process is used to hide a message in html document(css embed into html) and extracting process is used to extract the hidden message from the html documents.the message is hided into the pages using css embed into the html page.

Keywords: Stegnography,CSS,Markup Language ,Steganalysis,Information Hiding

I. INTRODUCTION

Steganography is the art (or science) of hiding sensitive information in such a way that its existence cannot even be proven [1]. The word steganography is derived from Ancient Greek and literally means "covered writing". Throughout history many different techniques have been used to hide messages, including – but in no way limited to – invisible inks, tattoos hidden under a head of hair, microdots and wax tablets. The digital steganography is divided into two separate subsections: techniques for hiding computer files in other computer files, and how a file system can be used to hide the existence of all the files stored on it. One of the ideas steganography tries to implement, especially with modern techniques, is plausible deniability: any act which leaves little or no evidence of wrongdoing or abuse. Using encryption on its own is not without its drawbacks: when a message is encrypted, the existence of the message can clearly be seen; whether this is the cipher text used hundreds of years ago, or the pseudorandom bits which are today's digital "cipher text". Being able to hide the encrypted message adds another layer of security. This can be as simple as a null cipher being used to hide a message encrypted with a Caesar-shift substitution cipher, or as complex as a file first being encrypted using the Blowfish [2] encryption algorithm, and then hidden inside a JPEG1 image using something akin to least significant bit insertion.Firstly, one must understand the components of a steganographic message. The "secret message" refers to the part of the message which is intended to be hidden. The "cover data" refers to the container in which the secret message is hidden. The "stego message" refers to the final product.From a top-down approach there exist three types of steganographic approaches:,Pure Steganography,Private Key Steganography,Public Key Steganography.These categories convey the level of security with which the stego message is embedded, transmitted and read [3].

II. HIDING TEXT USING MARKUP LANGUAGES

Considering methods of hiding information in XML/HTML documents, two main groups can be identified. The first group comprises techniques originating from the classical text steganography while the second group includes methods which make use of mark-up languages specific properties.The former group methods treat XML/HTML documents as text files and consist in embedding secret in a file by changing its content in a particular way, depending on information one wants to carry. The result can be achieved, e.g., by inserting additional white-spaces, making deliberate spelling errors or changing a font. Many different methods of text steganography exist and they are described in details in literature, e.g., [8-12]. Although they can be successfully applied to XML/HTML documents, they will not be analyzed further because they are easy to detect. Our intention is to analyze schemes of sending secret message, which do not affect the visible content of a web page presented by a web browser.It is an advantage of the techniques from the latter group, which exploit the ordered structure of the mark-up languages. Employing such techniques prevents the viewers from being alerted by unexpected changes of the content; in this regard, the visible web screen is entirely equivalent to the original document.The allowable modifications are constrained by the particular languages' specifications; see e.g. [13-15]. A number of approaches to HTML/XML steganography have been proposed. Now briefly presented advantages and drawbacks. One of the most important terms which allow comparison of different steganographic methods is a steganographic capacity which describes the amount of information, which can be hidden in a given covert channel. It is typically expressed as the maximum size of a secret message in bits. Obviously there is usually a trade-off between the capacity and the security, i.e. the bigger capacity a certain method offers, the easier it is to detect.

HTML Documents

Secret text can be easily concealed within HTML documents because HTML tags are case insensitive. For instance, the tags <a title="clients">, <a TITLE="clients">, and <a TitlE ="Center">, are all the same and have the same effect on the rendering of the document. Text steganography applied in HTML documents can be performed by changing the case of the letters that make up the HTML tags. In particular, the secret message is represented by the capital version of the tags' letters, or vice versa depending on the algorithm being used. As for the recovering process, all the capital letters from the HTML document have to be captured and concatenated together in order to produce the original covert message.

White-Spaces In Tags

In HTML we can add a white space before ">" sign, which marks the end of a tag. This way it is possible to send one bit of a secret message for each tag, by selecting the tags to which we want add an additional space before its end sign. Another possibility is inserting white-spaces after the tag (if there is no body text after it). Main drawback of this approach is that in consequence of its application to files, they are getting bigger, so the stego channel can be easily revealed.

Changing The Case of Letters In Tags

HTML tags are case insensitive, hence we can take advantage of it to hide a message within a document by changing the case of specific letters in a tag's name. For example, <ID>, <id>, <Id> and <ID> mean exactly the same and we can encode two bits by choosing one of its version. Big capacity is the main advantage of this method. On the other hand, it is very easy to discover the stego channel since it is very unusual to use small and capital letters alternately.

Using default values of attributes. Some of the HTML attributes have their default values defined. An HTML document is treated by an HTML viewer the same way whether the default values have been explicitly defined or not. It gives an opportunity to hide an additional information by specifying default values in some parts of the HTML document, and skipping them in other parts. This method is hard to discover, but the limited possible number of attributes having default values is its main drawback.Attribute order permutation. The HTML standard does not define a preferred order of attributes, what means that any order can be used without affecting a web page appearance. Since the order of attributes has no mean, this method can be applied without any restrictions. Changing the order of attributes for hiding information within an HTML document is the most interesting HTML steganography method. It does not change the original file size and it is hard to detect without computer programs analyzing the HTML document's structure.Having a tag comprising 8 attributes, there are 8! = 40320 different permutations, what lets one to hide over 15 bits of information within this single tag. This method is probably the most often mentioned one in the context of hiding data in the mark-up language documents. Its main advantage lies in its security, but in practice it allows only to send a small amount of data, because it is limited by the number of attributes being used in the original document.

III. ALGORITHM FOR ENCRYPTION AND EMBEDDING ENCRYPTED TEXT IN CSS:

Algorithm for decryption:

1. Convert the text into ASCII.
2. Calculate mod 26 of each ASCII value.
3. Store the quotient and remainder in separate variables.
4. Add 97 to each value of quotient and remainder to bring the values in ASCII range 97 to122 which is the range for small alphabets.
5. Convert quotient and remainder back to characters.
6. Generate the class selector by appending quotient to remainder and inserting 1between them to separate them.
7. Select a key and generate a single digit from it by successively adding its digit until we get a single digit.
8. The digit obtained will be the new position where the class selector will be inserted.
9. Now finally add the class selector to that position in CSS coding.

Algorithm for decryption:

1. First the class selector is extracted by the position given by key from CSS file.
2. Then both quotient and remainder parts of class selectors are separated as they are differentiated by 1.
3. Convert each of quotient and remainder into their ASCII value.
4. Now, subtract 97 from ASCII value of each of quotient and remainder.
5. Now, multiply each quotient value with 26 and add remainder to them.
6. Convert every value back to characters.

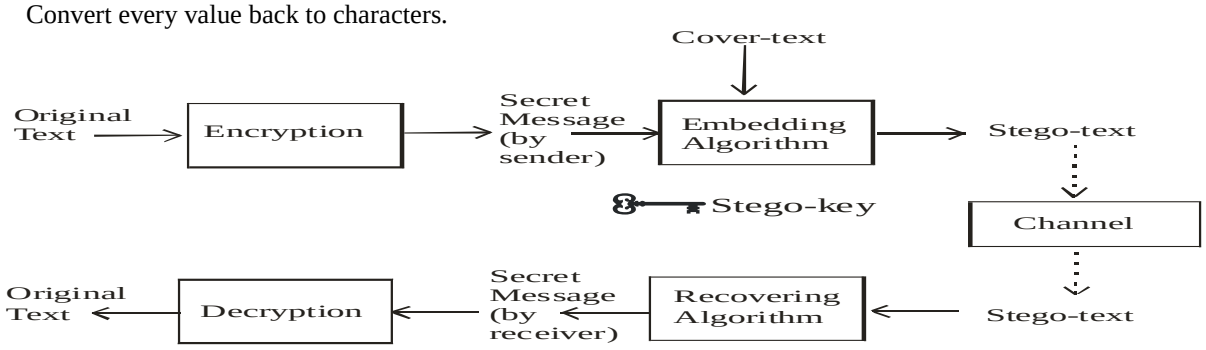

Figure 1: The Proposed Algorithm Structure

Motivation To Use CSS

The size of web pages is not noticed generally. So the data can be added to CSS coding of web page. Generally whenever we right click to view the source code then CSS code is not shown it makes the use of CSS imperceptible. In CSS coding of web page it is difficult to find a particular

class selector because there are various class selectors in CSS code from which some may be left undefined without being noticed The dynamic position of data in the CSS blocks the access to data. It motivates to hide the data in the CSS.

IV. RESULT IMAGES

The proposed work doesn't affect the web page as shown below. Firstly the original web page is shown The web page consists of data is shown.There is no difference in both pages.

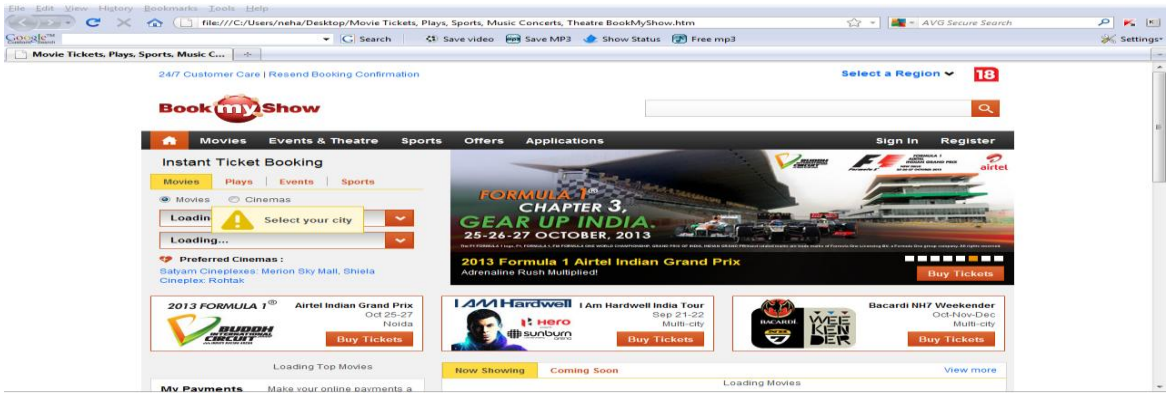

Figure 2: Web page before steganography

Figure 3: Web page after steganography

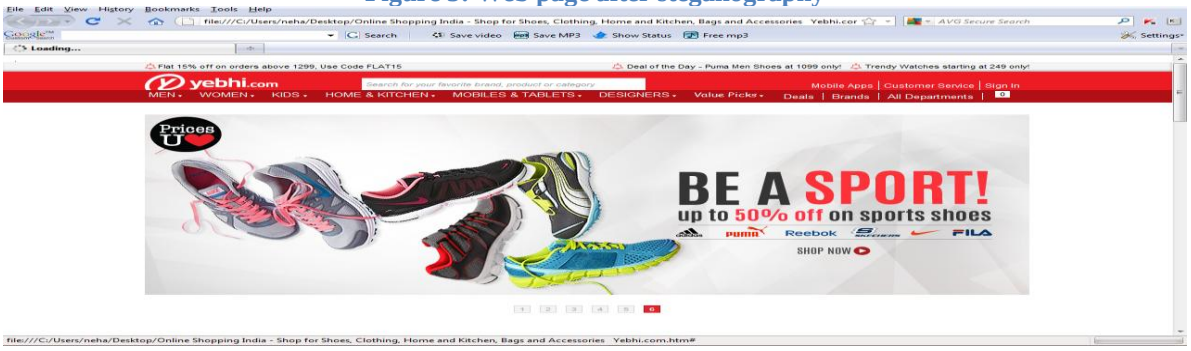

Figure 4: Web page before stegnography

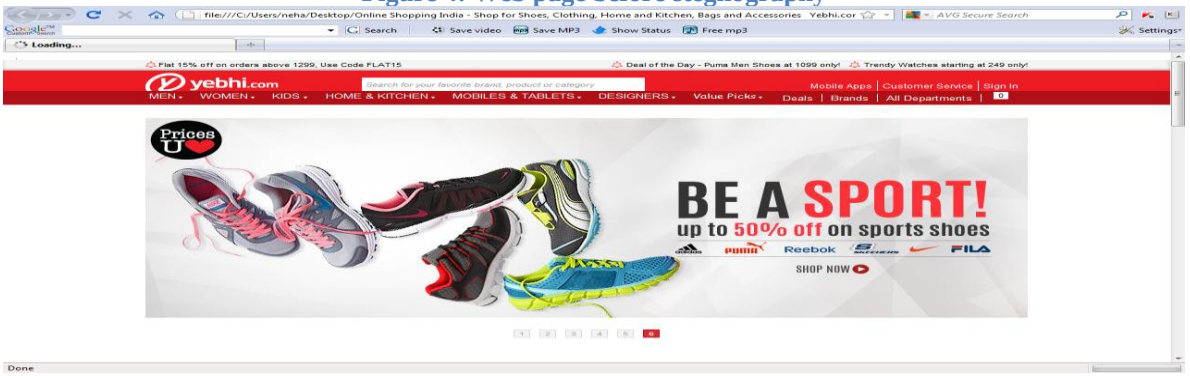

Figure 5: Web page after steganography

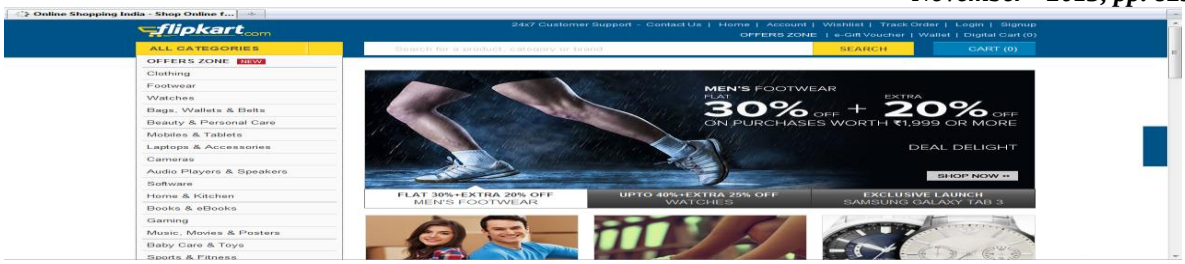

Figure 6: Web page before steganography

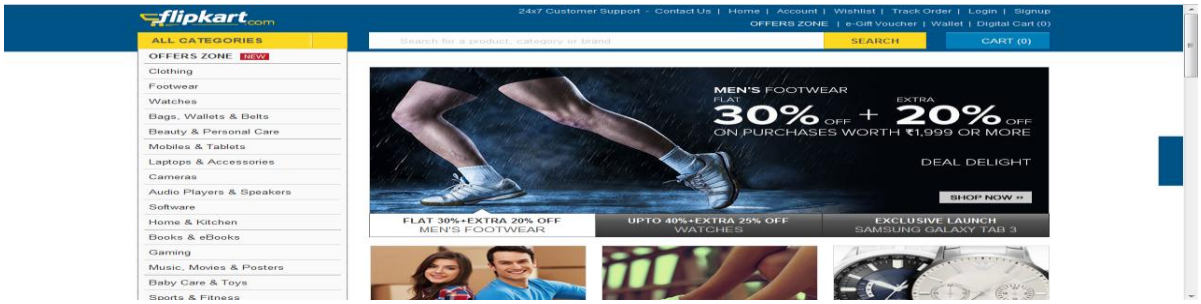

Figure 7: Web page after steganography

The new class selector which is generated from encrypted text has no visual impact on the appearance of web pages as shown in above figures. Figure: 3,5,7 are the original web pages before steganography and figure: 4,6,8 shows the modified CSS code of web page after insertion of class selector.

Table 1: Showing Analysis Results

Web Page	CSS file name	ACC Value (%)
Bookmyshow	Default	.034
Yebhi	Default1	.056
Flipcart	HPS-288cd996-nogz	.0053

Change in source of CSS can be analyzed by introducing a new parameter i.e. ACC (additional code changes).The value of ACC below the 10% is acceptable. Lower the value of ACC, difficult the data to be found. It means low value of ACC is better as it increases the security. ACC basically denotes the changes the source code of a file. It can be determined by comparing the source code of original and the stego file.So, ACC =sizeof code changes/size of original code *100;
It can also be given as Nc/No *100 where Nc is the number of characters changed and the No is the total of characters in original file.Here, in the default.css of bookmyshow.com is modified. The value of nc is 63493-63471=22,where 63493 are the number of characters of modified file and the 63471 are number oif characters ion original file. So, Acc=22/63471*100=.034%.

V. CONCLUSION

Steganography is to hide a secrete message within a cover-media in such a way that others cannot discern the presence of the hidden message. In simple words "steganography means hiding one piece of data within another". It is the major distinction between steganography and other method of hidden exchange of information. Example: in cryptography method people become aware of the existence of information by observing coded information, although they are unable to comprehend the information. Most steganography jobs have been performed on images, text, music, video clips and sound.Web based communication has a great amount of bandwidth and hence can be used for secret communication. Both HTML and CSS are two basic but important and universal tools for web development. New scheme on hiding information that is embedded through a Cascading Style Sheet (CSS) by using End Of Line (EOL) on each CSS style properties, exactly after a semi-colon.

REFERENCE:

[1]. Sch93 Bruce Schneier, *Fast Software Encryption*, 1993, *Cambridge Security Workshop Proceedings*, pp. 191-204, Springer-Verlag, 1994.
[2]. Yang, Li "Digital Watermarking". Canada, Ontario. University of Windsor, Novembe13, 2003.
[3]. Steganographic Techniques and their use in an Open-Systems Environment- Bret dunbar,The Information Security Reading Room, SANS Institute 2002 http://www.sans.org/reading-room/whitepapers/covert/677.php.

[4]. Anderson R.J. and Petitcolas F.A.P., "On the Limits of steganography," J. Selected Areas in Comm., vol. 16, no.4, 1998, pp. 474–481.
[5]. Arvind Kumar et. al., "Steganography- A Data Hiding Technique", International Journal of Computer Applications (0975 – 8887) Volume 9– No.7, November 2010.
[6]. Disappearing Cryptography: Being and Nothingness on the Net, Wayner Peter. 1996
[7]. Udit Budhiaa, Deepa Kundura. Digital video steganalysis exploiting collusion sensitivity.
[8]. F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, *Information Hiding. A Survey*, Proceedings of the IEEE, 87, 7, 1062–1078, July 1999.
[9]. T. G. Handel and M. T. Sandford II, *Hiding Data in the OSI Network Model, Information Hiding*, LNCS 1174, pp. 23–28, Springer Berlin, 1996.
[10] N. F. Johnson and S. Katzenbeisser, *A Survey of Steganographic Techniques, Information Hiding: Techniques for Steganography and Digital Watermarking*, pp. 43– 75, Artech House, 1999.
[11]. HTML standard specification, http://www.w3.org/TR/html4/ (last visited: 04.2010).
[12]. XML standard specification, http://www.w3.org/TR/REC-xml/ (last visited: 04.2010).
[13]. A. G. Memon, S. Khawaja, and A. Shah, *Steganography: A New Horizon for Safe Communication through XML*, Pakistan Journal of Theoretical and Applied Information Technology, 4, 3, 187–202, March 2008.
[14]. Monika Aggarwal et. al. "Text Steganographic Approaches: A Comparison" International Journal of Network Security & Its Applications (IJNSA), Vol.5, No.1,January 2013.
[15] Youssef Bassil et. al. , "A Generation-based Text Steganography Method using SQL Queries" International Journal of Computer Applications (0975 – 8887) Volume 57–No.12, November 2012.
[16]. Sabu M Thampi, "Information Hiding Techniques: A Tutorial Review" Thampi, Sabu M. "Information hiding techniques: A tutorial review." arXiv preprint arXiv:0802.3746,(2008).

