Theoretical and Applied Informatics
ISSN 1896-5334
Vol. 26 (2014), no. 3, 4
pp. 177 - 187

Network steganography method for user's identity confirmation in web applications

Paweł BUCHWALD[a], Maciej ROSTAŃSKI[a], Krystian MĄCZKA[a]

[a] University of Dabrowa Gornicza
ul. Cieplaka 1C, 41-300 Dabrowa Gornicza, Poland
{pbuchwald, mrostanski, kmaczka}@wsb.edu.pl

Abstract: Extensive use of computer networks is associated with the development of various effective methods that are suitable for hiding information in the contents transferred over the network. These methods are described as network steganography. Since web applications use HTTP protocol to transmit the requests to the server and send the answers to the final recipient, specifically HTTP protocol is ideal for hiding information. For example, there are several methods that can be used to transmit the additional content in the HTTP header. In this paper, we present authors' evaluation method for network steganography using HTTP specific properties and evaluate the effectiveness of some techniques, providing experimental results.
Keywords: HTTP, network steganography, web applications.

1. Introduction. Mechanisms of network steganography

Steganography is a set of methods to hide additional information in the publicly available data, in such a way that it is potentially unnoticed by everyone but the intended recipient. Steganography methods are often used to hide the content in digital media, such as images, videos, sound files, etc. One of the many uses of steganography can be copyright protection.

Steganography techniques involve hiding the additional information that fulfills the role of additional identifier. Such an identifier is capable of providing information about the owner or author of protected content [1]. One of the most popular steganography applications is hiding information by encoding data in the image. The sender is able to write additional data without noticeable loss of image quality by inserting this data into least significant bits of the image [2].

Extensive use of computer networks is associated with the development of various effective methods that are suitable for hiding information in the contents transferred over the network. These methods are described as network steganography (e.g. [3], [4]). The idea of network steganography involves the use of specific properties of

standard communications protocols that can be exploited to hide additional data. Network steganography methods enable the implementation of an additional communication channel, hidden in the open channel. Such data channel can be used to distribute additional meta data through the network during normal operation of the network. The advantage of network steganography methods used for the hidden data transmission is the lack of additional traffic being generated. Data transmission methods which are based on network steganography may utilize packet data units of several layers of communication. Data transmission method, which is based on network steganography, may use other protocols' data transmitted over the network, so the distribution channel data is dependent on the network traffic associated with the other programs and applications.

Network steganography reduces the potential receiver's ability to detect hidden information distributed through publicly available transmission channel, which is a big advantage of this method.

Methods of network steganography use of optional fields for standard protocols, use the redundancy of data, or interpretation of PDU units. For example, packets with incorrect CRC sum are rejected by the standard transmission equipment, but can be used by specially prepared software to interpret their content. Another way is to use some fields in the standard unit of encapsulation. One of such field is the Type Of Service (TOS) fields in Internet Protocol. This field is used in networks that support quality of service, but its interpretation is not strictly defined and depends on the specific implementation. In many networks, this field is not used, which gives the possibility of its application to transmit hidden metadata [5]. Another field that allows the transmission of additional data is the IP identification field. It must be unique for datagrams within a single transmission. This field has no defined relationship between numbers. Therefore, it is possible to construct an algorithm that uses this field to encode additional information [6].

Additional data can also be hidden in the fragmentation field, which is located in the header of the IP datagram. The most significant bit of this field is not used and can be applied as an information carrier in network steganography. IP datagram includes the option field that does not have a precise interpretation. It may have a variable length and in most cases is not used. Thanks to the free use of the field, we can use it to send additional metadata. This field is not modified by routers [7], which is an advantage when being used in the network steganography.

Also, TCP segments give the possibility to hide the distributed data. One option is to use the initial identification number field, which can be used once in a given session. This method is applicable when establishing multiple concurrent TCP connections. The advantage of this solution is difficult to detect because it requires monitoring of all the established connections. For the purposes of distribution of hidden information TCP Reserved field can also be used. Currently, these fields have no meaning in most networks, but in the future they may become important [8]. The disadvantage of steganographic solutions based on the TCP/IP protocol features is that they are relatively easy to be distorted by additional software. It is more difficult to analyze the

contents and removing additional hidden information for the protocols in the upper layers of the ISO / OSI. Such an analysis requires knowledge of the context of the transmitted information. One of the commonly used application layer protocols is an HTTP protocol. HTTP gives wide possibilities to hide additional information. If data are being transferred within the web application, the possibilities of hiding information are even greater, due to the less structured nature of the transmitted content.

2. Network steganography methods in web applications

Message Web applications use HTTP protocol to transmit the requests to the server and send the answers to the final recipient. HTTP protocol is ideal for hiding information. There are several methods that can be used to transmit the additional content in the HTTP header [9]. There are also solutions to the interpretation of the additional variables, that are sent directly in the URL to which the request was sent via the HTTP GET method. This method of hiding data can be used also in the business logic layer modules, based on the model of REST or Web Services. An example of a system using HTTP GET requests to send hidden data is hcovert system [10]. Blueprints of this system have been made publicly available on the website http://hcovert.sourceforge.net/ with open source code.

Significant number of websites using AJAX are based mostly on asynchronous JavaScript calls. This kind of communication with the server also provides wide possibilities for applications of steganography. The asynchronous nature of the messages sent to the server in solutions based on AJAX technology allows for more frequent exchange of data. One of the methods of language using JavaScript to hide the information is presented in publication [11].

In web applications that use standard HTML, server requests are issued only at particular queries of opening and refreshing WWW pages. Modern methods however utilize AJAX technology for invoking extra requests at the specified events (e.g. graphical interface related). Important fact is, this solution creates more transactions between web client and server. The more requests are generated in time, the more data can be transmitted with the means of a network steganography. The drawback lies in the character of a data transmission itself – the requests are asynchronically generated by callback functions. This means the time and frequency of such function calls are not deterministic values and this complicates the re-arrangement of transmission fragments of data in the steganografic channel. Additional mechanisms for transmission reliability are necessary.

An interesting method that allows sending hidden data is based on information contained in cookies [12]. Cookies are one of the most commonly used methods of data recording on the client (web browser) side. The syntax for creating a cookie includes many optionally set fields, and thus can be used to transmit additional information.

By specifying the appropriate type of content one can use the HTTP protocol to transmit graphic images. This allows to save the data directly in the image distributed via the HTTP protocol. The use of images is the most well-known method of steganography and often is based on the use of low bits – LSB method [13]. Proper interpretation of the least significant bits allows for the implementation of a communication channel using publicly available web services for the exchange of graphical content. One of these systems was presented in [14]. Widespread use of graphical content to send hidden data means that this hidden channel of distribution is vulnerable to discovery. In addition, there are tools for disruption of early-bits in images in order to eliminate steganographic message.

At present, widely accepted standard data storage format is XML. It is often used to represent the various types of data. His native form allows to save text content. It also allows to save images, as illustrated by SVG format. XML also allows you to record audio content (standard Music XML [15]). Extensible nature of XML markup allows for its wide range of applications. The importance of this type of recording format has increased with the introduction of HTML 5.0. Some of the properties markup languages such as HTML or XML may also be used to hide information - useful in this task could be the facts that:

- HTML is not case-sensitive;
- markup languages HTML and XML allow for bypassing whitespace;
- the order of attributes in the XML and HTML, which describes the tag is optional - it means that they do not affect the interpretation of tag;
- attributes of some of the HTML tags, such as ID and NAME, are optional and can be used for the purposes of saving the additional information.

An example of such a method of hiding information is presented in the next section of the article.

3. Effectiveness of hidden data channel based on HTML tags

Transmission efficiency realized by a hidden steganographic channel, which is based on HTML, can be defined as the ratio of the size of the data sent through this channel to the total size of the pages coded in HTML. Transmission efficiency can be expressed using the formula (1):

$$E = \frac{2^n}{s} \tag{1}$$

where.

- n - the number of bits used by the steganographic channel;
- s - the size of the HTML data used as a carrier (in bits).

Size of data that can be saved as a hidden in HTML tags depends on the steganographic method. Effectiveness analysis for covert channel was carried out, based on the following techniques that use the properties of HTML:

- Change the case of tags - HTML is not sensitive to the applied case. Changed case does not have any effect on page visualization (for example, tags
 and
 are interpreted by the HTML parser as the same), and can be used to encode additional information.
- Adding space characters at the end of tag - standard HTML parsers bypass whitespace characters such as spaces, tabs, and enter. This property can also be used to hide additional data. The tags "
" and "
" will be interpreted in the same way, and can be distinguished by additional software interpreting hidden transmission channel.
- Extending tags with additional attributes such as ID or NAME - These attributes are used in many cases to describe additional properties of tags. This method of hiding information is most effective, however it should be noted that these attributes are not present in all the tags. Their introduction does not change the appearance of the web page, however too frequent use of this method may lead to the disclosure of covert channel transmission. For this reason, it is assumed that the additional attributes should be added with a certain specified probability.

When these methods of hiding data in HTML pages are applied, the efficiency of the hidden transmission channel can be estimated using the formula (2):

$$E = \frac{n * (2 + P_w * C * L)}{8 * (\sum_{k=1}^{n} Z_n + D)} \tag{2}$$

The meaning of the symbols used in the formula is as follows:

- n - number of tags in the HTML page used for the implementation of the hidden transmission channel
- P_w - the likelihood of an additional attribute in the tag
- L - length of a string that specifies the value of the additional attribute
- C – number of hidden bits encoded in one character
- Z_n - the number of characters needed to encode the tag n
- D - the amount of additional data, such as the content of the page text, binary data, etc.

The presented method of analysis was used to evaluate the effectiveness of the transmission channel, and thus the size of the data that can be transferred through

hidden channel based on selected web pages online. The evaluation results are shown in Tab. 1.

Site	Tag count	Length [B]	Encoded data size [B]
www.wsb.edu.pl	700	32826	210
www.wp.pl	1729	103513	518,7
www.onet.pl	1805	102831	541,5
www.meteo.gig.eu	266	24538	79,8

Tab. 1. Effectivity of the transmission channel

4. Applications of steganographic methods and information security aspects

Many experts in the field of information security stresses out the problem of the possibility of using data hiding techniques for criminal purposes. Data redundancy and optional fields in network data formats allow for the implementation of additional hidden information channel, so it can be used to transfer information between criminal groups, whose detection can be very difficult [16].

Steganography can also be used in the law enforcing and computer systems security improvement. An example of this might be digital content marking, such as images or music, with special steganographic tag. It allows for unambiguous copyright information placement [17].

Steganographic methods are also used in improving security of publicly available wireless networks based on the 802.11 standard [18]. They can be used to implement additional transmission channel, which would be used to authenticate the access point. An interesting way to implement covert channel is leveraging Timestamp field in Beacon frames, which are periodically sent by the access points. The proposed channel is unidirectional and can be designed to perform authentication of the access point, thereby reducing the man in the middle attack possibility. Hidden transmission channel can be realized in wireless networks IEEE 802.11 standard by intentionally marking frames as defective or the use of fields to fill PAD physical layer frames.

The concept of the use of covert channels of communication can be used to confirm the identity of web applications users. Confirming the identity of computer systems users through redundant transmission channels, is used in many information systems such as electronic banking systems. A popular method of authentication is a one-time code transmission through the GSM network [19, 20]. An example of a user authentication system is the solution working with an online passport Axionics that uses redundant video channel for transmitting authentication information. Video channel is realized by the blinking code on the computer screen. The code is read by a dedicated device and then converted to a single token entered by the user in a web

form. The remote server verifies the authentication token and thus verifies the user's identity. In both methods user interaction is required, by using the web form. The use of steganographic mechanisms to implement additional authentication data transmission channel, allows for automation of the identity of the user confirmation process. In this embodiment, the authentication data is transmitted using a hidden channel data in HTTP. Solemn use of the mechanisms of steganography does not guarantee the security of data transmission, because there is some probability of discovery of the steganographic channel. Therefore, it is necessary to apply additional encryption mechanisms, such as asymmetric encryption. The use of network steganography allows for managing user credentials and authentication in an easier way, because the channel's data are associated with each other. Concept of such automated authentication system is illustrated in Fig. 1.

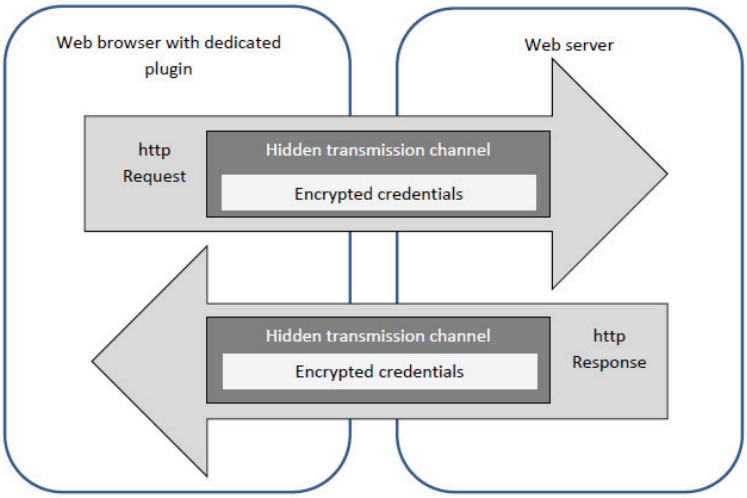

Fig. 1. Authentication with covert channel transmission

This method of user authentication involves the use of a dedicated Web browser that encrypts a string of one-time authentication using public key. Encrypted authentication string is sent via hidden communication channel with steganography methods previously presented. On the server side it is possible to interpret the data sent and decrypted with one-time authentication string. The server generates its single string authentication, joins decrypted authentication string received from the client and then encrypts and sends the data via the transmission channel hidden in the HTTP response to the client browser. The client application reads data, decrypts it, and is able to check whether a received one-time authentication string is identical to the one previously sent. Decrypted authentication string generated by the server is merged with another one–time authentication string created on the client side, and the

encrypted message is sent to the server. The server repeats the previously described operation by checking whether the received authentication key is identical to the previously sent. In this scenario, one can create each response and authentication requests that are received or sent by the server or the client's browser. The presented method enables the use of security RSA asymmetric encryption algorithms for confirmation of the identity of the user. This method also creates the possibility of using steganographic channel, which is formed during HTTP transmission. Therefore, additional synchronization between the data channel, and the authentication channel is not required. The solution is an additional mechanism for authentication and security improvement. It requires a dedicated client application, which reduces its flexibility. It should also be noted that the proposed method is dependent on steganographic channel.

4. Conclusions and summary

In this paper, we present authors' method for network steganography using HTTP specific properties and evaluate its effectiveness. Network steganography is a method of hiding data transmission channel and next to cryptography is an interest to specialists in the field of information security. The paper shows that the method of steganography can also be used as a tool for improving the safety aspects of data transmission systems. The presented example shows how to confirm the credentials in web applications using hidden information channel. This method can be used together with other mechanisms of user verification systems that can contribute to improving security [3]. The proposed method does not need to be limited only to solutions for Web applications, but can be used in other application layer protocols for the implementation of covert channel transmission.

Analysis of the efficiency covert channel data acquisition has shown that it can be used to distribute some metadata that are associated with an open channel of transmission. Hidden bandwidth transmission channel is dependent on the structure of the open channel of the distributed data, which are used as the hidden data carrier. In this article, authors propose novel methods to hide data in HTML tags. The extension of this solution with additional coding possibilities, such as by using a sequence of attributes of HTML tags or additional attributes such as HTML styles would allow for the bandwidth covert channel increase. The presented methods allow for the implementation of the transmission channel that has parameters for two-way communication and can be used for sending of identity confirming codes. This solution can automate the process of confirming a user's identity by eliminating the need of having user to manually enter one-time passwords via a web form. Authentication method implemented using steganographic channel allows to hide the process of confirming the identity from unauthorized users and can be used to differentiate web server responses, depending on whether the user's identity has been confirmed.

References

[1] Koscielny C.: Steganografia. Biuletyn Politechniki Zielonogórskiej, grudzien 1999. online: http://www.pz.zgora.pl/pz/biuletyn/grudzien99/b14.pdf.

[2] Farid F.: Detecting steganographic messages in digital images. Institute report, Dartmouth College, Computer Science, 2001. online: http://www.cs.dartmouth.edu/~farid/publications/tr01.html.

[3] Polak L., Kotulski Z.: Sending hidden data through www pages: detection and prevention, in: Engineering transactions 58, 1–2, 75–89, Polish Academy of Sciences 2010

[4] Szczypiorski K., Mazurczyk W.: Steganography in IEEE 802.11 OFDM Symbols, in: International Journal of Security and Communication Networks Vol. 3:1-12, ISSN: 1939-0114, John Wiley & Sons 2011

[5] RFC 1349, Type of Service in the Internet Protocol Suite.

[6] RFC 791, Internet protocol DARPA Internet Program Protocol Specification September 1981

[7] Rostanski M.: Protokół IPv6 jako następca IPv4 w sieciach przedsiębiorstw. Ciągłość działania systemów migrowanych do IPv6, Wyższa Szkoła Biznesu w Dąbrowie Górniczej, ISBN 978-83-62897-77-3, Dąbrowa Górnicza, 2014

[8] RFC 793, Transmission Control Protocol, DARPA Internet Program, Protocol Specification September 1981.

[9] Dyatlov A. and Castro S.: Exploitation of Data Streams Authorized by a Network Access Control System for Arbitrary Data Transfers: Tunneling and Covert Channels over the HTTP Protocol, tech. rep., Gray-World, June 2003,

[10] Hcovert software homepage: http://hcovert.sourceforge.net/

[11] Bauer M.: New Covert Channels in HTTP: Adding Unwitting Web Browsers to Anonymity Sets, in proc: Privacy Electronic Society, Oct. 2003,

[12] Castro S. and Gray World Team, Cooking Channels, hakin9 Magazine (www.hakin9.org), May 2006, pp. 50–57

[13] Van Schyndel, R., a.Z. Tirkel, , Osborne, C.: A digital watermark. In: Proceedings of the IEEE International Conference on Image Processing. IEEE Comput. Soc. Press; volume 2; 1994

[14] Burnett, S., Feamster, N., Vempala, S.: Chipping away at censorship rewalls with user-generated content. In: Proceedings of the 19th USENIX Conference on Security. USENIX Association; USENIX Security'10; 2010.

[15] Lombardo V. et al.: The virtual electronic poem (vep) project. In: Proceedings of the 2005 International Computer Music Conference. 2005. p. 451-454

[16] Grochowski L., Hołdys B.: Steganografia a zagrożenia cyberterrorystyczne, w: Prokuratora i Prawo 7-8, 2013r online: http://www.ies.krakow.pl/wydawnictwo/prokuratura/pdf/2013/07-08/

[17] Urbanovich N., Plaskovitsky V.: The use of steganographic techniques for protection of intellectual property rights. In: New Electrical and Electronic Technologies and their Industrial Implementation (2012), p. 147

[18] Tolani R., Yeole A., Gavhane S.: An HOTP Based Algorithm to Enhance Wi-Fi Security, in: International Journal of Advanced Research in Computer Science and Software Engineering; July 2013 ISSN: 2277-128X

[19] Buchwald P.: Użycie redundantnych kanałów informacyjnych do poprawy bezpieczeństwa portali internetowych w: Grzywak A. (red.): Internet w Społeczeństwie Informacyjnym - Zastosowania Internetu, WSB Dąbrowa Górnicza 2009

[20] Grzywak A., Klamka J., Buchwald P., Pikiewicz P., Rostanski M., Sobota M.: Podpis elektroniczny i identyfikacja użytkowników w sieci Internet, WSB Dąbrowa Górnicza 2013

Metoda steganografii sieciowej dla potwierdzania tożsamości użytkownika w aplikacjach webowych

Streszczenie

Szerokie zastosowanie sieci komputerowej jest związane z rozwojem wielu skutecznych metod, które są odpowiednie do ukrywania informacji w treściach przesyłanych poprzez sieć. Metody te są określane mianem steganografii sieciowej. Ponieważ aplikacje internetowe używają protokołu HTTP do przesyłania żądań do serwera i wysyłania odpowiedzi do końcowego odbiorcy, protokół HTTP jest idealny do ukrywania informacji w szczególności. Na przykład, istnieje kilka metod, które mogą być wykorzystane do przesyłania zawartości dodatkowych w nagłówku HTTP. W tym artykule przedstawiono autorski sposób oceny metod steganografii sieciowej za pomocą konkretnych właściwości HTTP i dokonano oceny skuteczności niektórych technik, podając wyniki doświadczalne.