DESIGN OF STEGANOGRAPHIC ALGORITHM USING WEB SERVICES

Liberios VOKOROKOS, Ján HURTUK, Branislav MADOŠ, Kamil FRIGA
Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic, tel. +421 55 602 3023,
e-mail: {liberios.vokorokos, jan.hutuk, branislav.mados}@tuke.sk, kamil.friga@student.tuke.sk

ABSTRACT

The aim of the paper is focused on the design of steganographic algorithm that is using web services as the cover medium. The principle of steganographic algorithm is to insert a secret message to an ordinary communication, which in the background with the help of an automated system generates this secret message to a desired recipient. Covert communication is available to public and it is not hidden from third parties. This paper summarizes the results of the algorithm and at the very end there is also testing of the software application that was used as the proof of concept.

Keywords: *steganography, social network, covert medium*

1. INTRODUCTION

Designing of applications and information systems into several areas such as stores of various goods or scientific research is an unstoppable trend these days. Along with this phenomenon goes hand by hand also a technology of secrecy, hiding and concealment of various sensitive data as well as maintaining certain anonymity in the world of internet technologies.

There are several forms of transmitted or sent data protection accompanied to. One of the possible ways of avoiding of communication trapping with the assistance of third sides is steganography.

1.1. Steganography

Steganography, originally composed from Greek words „steganos“ – hidden and „graphein“ – writing is a discipline which is being studied and used for thousands years. Expression in the form of steganography was given at the turn of the 15th and 16th century. The beginning of using this expression in IT sector was in 1991.

The main principle of steganography is hiding of secret information in digital data, so the third side (unauthorized person) would be unable to detect the existence of the information.

The first step in creating of steganography model is choosing of cover medium, which is a secret message carrier. Any multimedia can form a cover medium, for example text, static images, audio, video etc. But there is a principle involved, that this media must contain a sufficiently large amount of redundancy, which the secret information needs to be inserted in.

The second step needs to include selection of appropriate algorithm for data hiding into the cover medium. This algorithm must be exactly the same also for the receiver. The secret information is inserted into the cover medium via algorithm, and this process creates a stego object (stego medium). For demand of higher security, some algorithms apply a secret key in process of hidden information inserting [1].

- **Secret message:** The secret message or information to hide.
- **Cover file digital medium**: The data or medium which conceals the secret message.
- **Stego file:** A modified version of cover that contains the secret message.
- **Key:** Additional secret data that is needed for the process of embedding and extracting of the secret message.

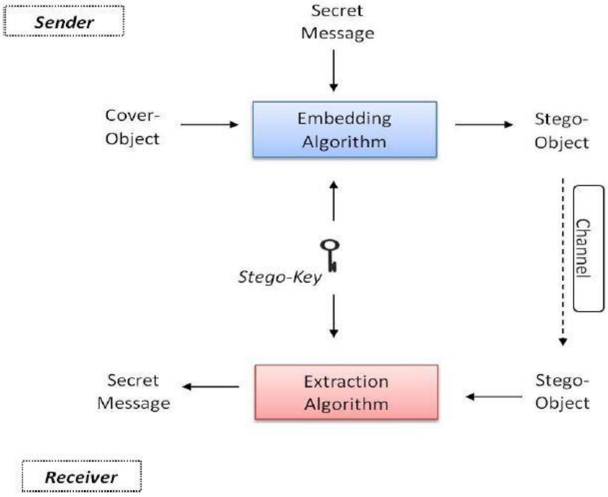

Fig. 1 Steganographic process of sending and receiving of stego message [1]

In the side of receiver, who is familiar with steganographic algorithm, there is being extracted a secret message from the stego object using a secret key, which was used during the process of inserting secret information.

Just as against every weapon, there is a counter-weapon and against any virus, there is a vaccine, for a modern steganography there is also no exception. The steganography is the part of steganology, and is aimed to bring algorithms for hiding the information. There is also steganalysis as the part of steganology, which is aimed to allow finding of secret messages that are concealed by steganography.

Steganographic system is characterized by the following features [1]:

- Undetectability
- Capacity
- Perceptual transparency
- Security and robusteness

Undetectability is the basic feature that tells us the number of performed statistical changes, caused by inserting a secret message. Capacity is the expression of maximum amount of secret information that can be inserted into the cover medium by using this method. Capacity is measured in bits. The security of the system is an attribute that protects a secret message, or expresses difficulty in extracting a secret message from designated medium. Robustness of the system is the resilience of the system, considering the attacks, mutilation or attempts to remove a secret message from designated media. The perceptive transparency is an ability of maintaining the required quality of the cover medium even after inserting secret information. The observer is not able to perceive a lower quality, respectively to recognize the cover medium from stego medium. This method is required especially for audiovisual data [1].

1.2. Analysis

Hiding information in text is the most important in steganography. One of methods was to hide a secret message in every nth letter of every word of a text message. After booming of Internet and different types of digital file formats it has decreased in importance. Text stenography using digital files is not used very often because the text files have a very small amount of redundant data [1].

When we are analysing the hidden communications solutions in selected web services, it is necessary to find current awareness of a number of respondents in the area of information technology. The testing was performed on 130 people, who have received information about steganography together with attached questionnaire. This sample told us, that more than a half of respondents, precisely 64% was not familiar with this concept and never heard about it, but the enormous interest in the possibility of testing or trying out one of the options resonated with greater range. Each and every one of these respondents responded to an additional question, that he would like to keep his internet communication only for authorized people.

The main mission is to design and subsequently implement the application, which seeks to interact with users based on the classic forms of messaging through the web interface. The essence of future applications is to send classified information from the initial informant to the end user, even when given unsolicited communications monitored by a third party. The aim is to create a new principle, and the protection of data provided.

2. PROPOSED ARCHITECTURE

For more specific understanding is required to fill a typical interaction between users and applications based on MVC architecture used in this application.

1. The user in the user interface performs an input action.
2. This action identifies Controller.
3. Subsequently, the controller decides what will follow respectively. What event occurs.
4. Controller also changes the value in the model or affect the view.
5. Then there will be updates View and change user are displayed in the model there is reportedly sending out chat messages [2].

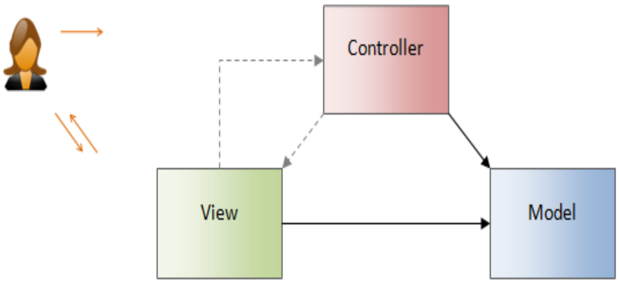

Fig. 2 Model View Controller (MVC)

The created algorithm is using a common communication between two users of social network. As it was mentioned above, occurrence of the communication is not hidden from third parties.

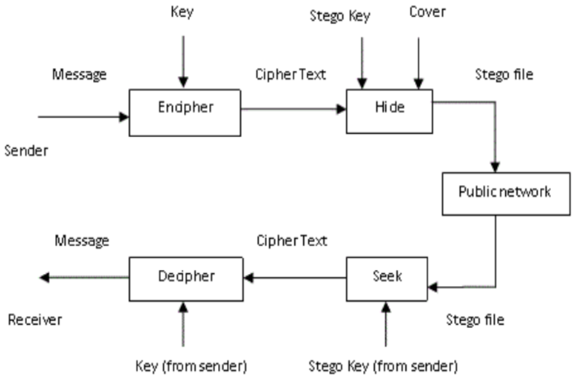

Fig. 3 Steganography with use of web services

2.1. Algorithm and work with binary

The administrator of conversation and the destination address knows the key that is based on the start of the process with stenography algorithm. The initial step in identifying steganography communication is defining the text of a confidential message to the form element that is transformed into a binary text string.

For Example, letter „Š“, acquires the binary systems value as 10001. Subsequently, there comes an addition of non-confidential report, which is not important and relevant and is sent along with a confidential message. PHP processes the binary code and applies it via CRON by dividing the input text string at certain parts. Consequently, there is a sending of messages in the chat by the process of chronological distribution.

Input Secret character, respectively letter = at least 16

The format of the binary system

0 marks the specified interval N, `<i = 1, j = 5>` seconds between two sent messages.

1 marks the specified interval N, `<i=6,j=10>` seconds between two sent messages.

The resulting report generated through CRON

```
15:31:20 Lorem ipsum dolor sit amet
15:31:30 consectetur adipiscing elit
15:31:31 Nam eu ornare
15:31:34 Nulla. Morbi luctus nec massa non imperdiet.
15:31:35 Integer consectetur malesuada.
15:31:39 Aliquam ut nisi
15:31:42 Vitae mi tincidunt
15:31:45 Auctor et id diam
15:31:47 Maecenas interdum massa
15:31:57 Orem, sit amet congu.
```

Result

- Letter Š = 10001010
- The number of rows = 10
- Time conversation range = 37 seconds

Due to security and a further reduction of noise there comes to expansion in terms of increasing the interval for values 0 and 1 after every 3 minutes of communication.

`N = <ix2,jx2>`

Reduction and higher effectiveness of solution
In case of duplicate appearance of values 0 or 1. After the split of the chain takes place delay, so 4x3 = 12 seconds.

The resulting report generated through CRON

```
15:31:20 Lorem ipsum dolor sit amet consectetur adipiscing elit
15:31:30 Nam eu ornare nulla. Morbi luctus nec massa non imperdiet
15:31:31 Integer consectetur malesuada Aliquam ut nisi
15:31:34 Vitae mi tincidunt
15:31:46 Auctor et id diam Maecenas interdum
15:31:55 Massa orem, sit amet congu
```

Result

- Letter Š = 10001010
- The number of rows = 12
- Time conversation range = 59 seconds

For casual sending of a telephone number we need around 9-10 minutes of conversation and from 110 to 120 lines of conversation. Regular addressing of Alejova 2 takes about 7-8 minutes of conversation and 79 to 92 lines of conversation.

2.2. Implementation

Because it is a dynamic Web application that does not have to be installed on the client computer for the reason of starting it with a web browser and its start depends on the server, is the solution implemented on a local server and then deployed into the production environment. The project is built on the backbone of the Apache, PHP and MySQL.

Server parameters:

- PHP in version 5, 5.3 and 5.4
- Encrypted access via a secure hypertext transfer protocol HTTPS
- CRON support for automatic execution of scripts
- Custom setting option modification PHP
- .htaccess
- MySQL database system
- Web Acceleration PHP Cache, Varnisch, Memcache
- Regular data backup
- Logs error and webhosting namely input

3. TESTS

Testing of the proposed algorithm was divided into two main parts:

- Testing of the server side
- Functional measuring of the system

3.1. Server testing

Considering the potential of the project realization for three local servers, application has been tested on different platforms, specifically it was a Windows server in version 2008 and 2012 and Gentoo Linux distributions, where everything took place in order and deviations were in terms of speed in a differential intervals not more than 3 seconds. This parameter also depends on the speed of Internet connection of users. From the server point of view there is no difference of more than 1 second.

3.2. Functional measurement system

Speed functionality measurement of the system, respectively individual modules loading was executed through PHP functionality. The test was conducted 3 times for each function, and the values were averaged.

The values with the response using the normal personal computer and Internet connection speed with a number of about 46,000 Kbps been measured and written in this table.

Table 1 Tests of functions

Function	Time
Login user	432
Logout user	395
Setting steganography	511
Send message	495
Receive message	321

3.3. Other tests

Thanks to the W3C Validators, elements in optimization for web browsers have been modified. This page is displayed without problems and almost identical to some deviations caused by HTML5. The application has been tested with Opera beta, Firefox 34.0, IE 11 and Google Chrome 42.0.

4. CONCLUSIONS

By this project, there is an option of creating the possibility of effective and protected communication with the purpose of concealing information regardless of capturing or monitoring by third sides. Common types are effective but more and more addressed to these types of steganography analytic tools to identify stegomedia. The main benefit is to protect vulnerable data, for example system deployment to a company as an additional protection against industrial espionage and especially where it is not possible to make free use of cryptography. The advantage is simplicity and speed of interactive communication, non-problematic check-in. The possibility of setting the principle of leadership communication is a unique possibility to create your own principle of communicating procedures for certain possibility of custom developing solution or research opportunities respectively. Visualization of web application is adapted for multiple devices such as mobile phones or smartphones to a standard PC which is also an advantage for the growing trend of using smartphones.

Steganography is relatively a generational discipline. This classifying messages principle has influenced the ancient sovereigns as well as world wars or Christianity. Despite non-characteristic scientific principles it is increasingly used in terms of formal planes 21st century and it is also the implementation of keys used by steganography algorithms as well as security and capacity perspective of covers. In terms of informal plane, there is classified a characteristic feature and that is a typical application of new original ideas based on free creativity. It does not require placing a strict emphasis on computerization of theoretical or mathematical principles.

In the future, there is a possibility of the extension of the system with additional add-on modules by constantly changing requirements either in inventing a new form of communication or information technology developments as well as the principles of programming languages.

Another revolutionary scientific step in this discipline in the near future does not impose a positive response and this may be the impulse for free individual research opportunities, or further development. Moving ideas of this communication concept are based mainly on the new creative stimuli. Aspects of modern steganography in computer science calculate with the possible implementations in secret counting located in an operating system environment that is fully in response of the striker, respectively the analyst as well as streamlining the methods of destruction of stop embedding without any possibility of disruption of a clearly visible value of the object. New statistical properties of various types of data objects can be also helping.

ACKNOWLEDGMENTS

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and KEGA 008TUKE-4/2013 Microlearning environment for education of information security specialists. The projects are being solved at the Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice.

REFERENCES

[1] KUNNEMANN, R. – Planning a Jailbreak: Use Steganography [online]. 2007. [cit. 2015-03-02]. Dostupné na internete: <https://www.infsec.cs.uni-saarland.de/teaching/SS07/Proseminar/slides/kuennemann-stego.pdf>.

[2] WATERSON, K.: Model View Controller MVC [online]. 2015. [cit. 2015-04-02]. Dostupné na internete: <http://www.phpro.org/tutorials/Model-View-Controller-MVC.html>

[3] BLOISI, D. D. – IOCCHI, L.: "Image based Steganography and Cryptography", In VISAPP07, pp. 127-134, 2007.

[4] RADHAKRISHNAN, R. – KHARRAZI, M.– MEMON, N. – Data Masking: A New Approach forSteganography, The Journal of VLSI Signal Processing, Volume 41, Number 3, November 2005.

Received July 30, 2015 , accepted August 28, 2015

BIOGRAPHIES

Liberios Vokorokos (prof., Ing., PhD.) was born on 17. November 1966 in Greece. In 1991 he graduated (MSc.) with honours at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He defended his PhD. in the field of programming device and systems in 2000; his thesis title was "Diagnosis of compound systems using the Data Flow applications". He was appointed professor for Computers Science and Informatics in 2005. Since 1995 he is working as an educationist at the Department of Computers and Informatics. His scientific research focuses on parallel computers of the Data Flow type. He also investigates the questions related to the diagnostics of complex systems. He is a dean of the Faculty of Electrical Engineering and Informatics at the Technical

University of Košice. His other professional interests include the membership on the Advisory Committee for Informatization at the faculty and Advisory Board for the Development and Informatization at Technical University of Košice.

Ján Hurtuk (Ing.) was born on 4th October 1988 in Kežmarok, Slovakia. In 2013 he graduated (MSc.) at the Department of Computers and Informatics at the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. Since 2014 he is studying as a PhD. student at the Department of Computers and Informatics at the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. His scientific research is mainly focused on the computer security.

Branislav Madoš (Ing., PhD.) was born on 20th May 1976 in Trebišov, Slovakia. In 2006 he graduated (Ing.) at the Department of Computers and Informatics at the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. He defended his PhD. in the field of Computers and computer systems in 2009; his thesis title was "Specialized architecture of data flow computer". Since 2010 he is working as an Assistant Professor at the Department of Computers and Informatics. His scientific research is focused on the parallel computer architectures and architectures of computers with data driven computational model.