CryptoCloak as a Protection Against Internet Surveillance

Dijana Vukovic
Department of Telematics
Norwegian University of Science and Technology
Trondheim, Norway
dijanav@item.ntnu.no

Abstract — ***CryptoCloak is an open source project that uses cryptographic algorithms for key exchange and encryption selected by the cryptographic community as solid and secure algorithms BUT it is doing the encryption in a clandestine manner. In the light of the latest Internet spying scandals, the goal of the project is to diminish the success of the automatic traffic analysing engines of anyone who has the access to the chat servers. Those analysing engines will notice just a cheap chat conversation, while the real encrypted information will be incorporated deeply in that cheap chat. The cheap chat is the cloak to the encrypted information communicated between two CryptoCloak clients. The first realization of the CryptoCloak uses the Skype API for Java programming language.***

Keywords - NSA surveillance, Diffie Hellman key exchange

I. INTRODUCTION

After *The Guardian* published news about the NSA collecting phone records [1] whole world was in shock. This news has caused the avalanche effect: more and more information about the NSA spying on private conversations (phone calls, chats etc.) start appearing. *The Guardian* kept on publishing articles related to this topic. The next step was an article about PRISM [2]. The whole timeline of information discovered about NSA surveillance can be found here [3].

As a response to this violation of privacy, Bruce Schneier wrote the article "*The US government has betrayed the internet. We need to take it back.*" [4]. The main idea of this article was that the engineering community need to bring back Internet as it used to be to the people - "*This is not the internet the world needs, or the internet its creators envisioned. We need to take it back. And by we, I mean the engineering community. Yes, this is primarily a political problem, a policy matter that requires political intervention. But this is also an engineering problem, and there are several things engineers can – and should – do.*" Guided by this idea, we started the *CryptoCloak* project.

The basic idea of the CryptoCloak can be described as the following: we will use the solid and secure algorithms that have been proven as secure in the last 30 years, BUT do the encryption in a clandestine manner. The automatic filtering rules of the spying agencies will notice just a cheap chat conversations, while the real encrypted information will be incorporated deeply in that cheap chat. The cheap chat will be our cloak to encrypted information. In the first phase of the project we used the Skype API [5] for the Java programming language to simulate cheap chat communication over the network. The next step will be implementation of a Java open source chat client for the most popular web browsers.

According to [6], computer scientists are involved in enabling intrusion on individual privacy and this lead to breaking the code of ethics [7]. This project will be our attempt to follow the basic idea of [4] and to enable use of the code of ethics in the right way – to contribute to society and human well-being without doing harm to others, and with respect the privacy of others.

II. SURVEILLANCE AND PRIVACY

Surveillance can be defined as "*close observation of a person or group, especially one under suspicion*". Many terroristic activities during past years lead to electronic surveillance. Law enforcement agencies needed the ability to conduct electronic surveillance to prevent crime, terrorism or any kind of malicious activities exploiting the Internet. Many people have opposed surveillance because it can be considered as an invasion of privacy (as with hidden video cameras) or a tool of social control (as in monitoring workers) [10].

To avoid issues of using surveillance, perpetrators of surveillance use five methods to minimize adverse actions to their actions [10]:

1. Cover-up and exposure: surveillance is commonly carried out as a secret – when people are not aware of it, they will not be concerned about it.
2. Devaluation and validation: expression "if you have nothing to hide, you have nothing to fear" can lead to a conclusion that surveillance is justified – if you are concerned about it, you must have something to hide.

3. Interpretation struggles: proponents of surveillance usually provide a believable justification for measures, or simply assume their effectiveness.
4. Official channels: experts of court are used to give an appearance of justice.
5. Intimidation, bribery, and resistance: surveillance measures can be intimidating – no one likes to think about the fact that their conversation will be recorded. Individual resistance to surveillance can appear in different shapes: avoiding detection, refusing to provide information, and encouraging surveillance agents not to enforce regulations.

Surveillance can be justified in some cases, as a cracking down a crime, or increasing efficiency of service systems, but it can also be a big threat to privacy.

Privacy can simply be defined as “the right to be left alone”. Privacy can be explained using two paradigms [11]: privacy is secrecy (confidentiality) and privacy is control. The second paradigm – privacy is viewed as the effect of individual control – can include several recommendations: individuals should be aware of how their information is collected, used, transferred, and retained, individuals have right to consent to it, individuals should have access to their information, and permission to correct them.

Privacy is the right of each individual, and it should not be threatened if it is not harmful to the others. These surveillance and privacy issues were the main reason for the information security community to start developing solutions for their solving. The CryptoCloak project started with the same intention.

III. THE CRYPTOCLOAK PROJECT

Spying engines do the traffic analysis as it is shown in Figure 1.

Figure 1. Traffic analysis

The first case is to filter the content looking for particular keywords, e.g. bomb, terrorism, etc. This content will be analysed to prevent potential terroristic attacks or similar issues. In the second case, any encrypted content will be stored for further analysis. Third case - noticing cheap chat conversations, e.g. "Hello!", "How are you?" - will be ignored. That is the fact the CryptoCloak uses to do the Diffie Hellman key exchange in a clandestine manner.

The CryptoCloak project is now in a testing phase. The implementation is done using Java programming language with additional API for Skype [5]. Skype was chosen as one of the most popular instant messengers worldwide. The way the CryptoCloak actually works is shown in Figure 2.

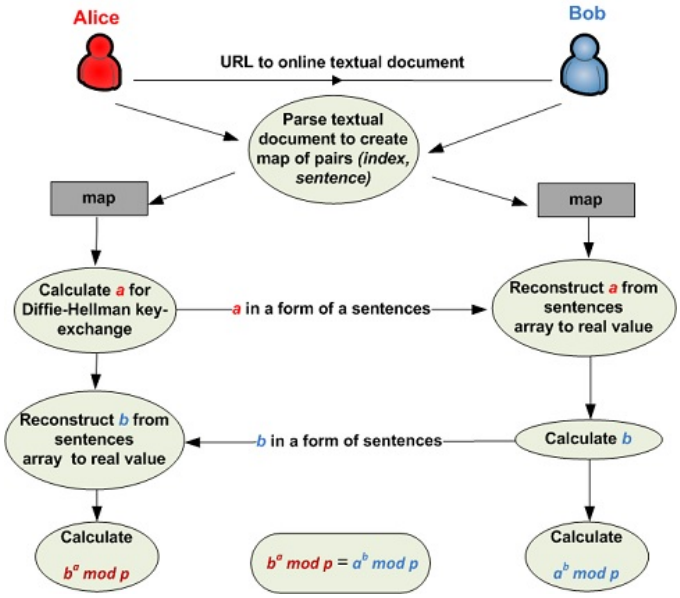

Figure 2. The CryptoCloak - Diffie Hellman key exchange

The Diffie Hellman key exchange is done in the modified way - instead of regular primes the sequence of sentences is going to be sent. The first step was finding free online available text files to provide the same source of sentences for both sides in communication (Alice and Bob) for encryption and decryption. The current idea is using free the *e-books* published under the "Project Guttenberg"[1]. There can be found over 40000 *e-books* from different genre. Alice chooses one *e-book* in textual format from the "Project Guttenberg" official web site and copies its URL, e.g. "The Jungle Book" - http://www.gutenberg.org/cache/epub/35997/pg35997.txt. She sends this URL to Bob. On both sides, using the same algorithm, the chosen text file is transformed into a map of pairs *(index, sentence)*. Only different sentences will come into consideration during the transformation. The next step is calculating the number *n* that fulfils the condition $2^n < sn < 2^{n+1}$, where *sn* is the number of sentences in the set after the transformation. For "The Jungle Book" *sn*=2121, which gives us *n*=11.

Using Diffie Hellman key exchange algorithm, *a* should be sent over the network from Alice's side. *p* and *g* are public, and they are built into the application. For the first phase of implementation *p* was 1024-bits prime. The *p* length is fixed to 1024 bits, but the *a* length may differ. This length is also sent as a sentence - sentence is chosen from the map where the index value is equal to the length value. After calculation, *a* is transformed into its binary value, then split into blocks of *n* bits. For every block, decimal value is recalculated, and used

[1] http://www.gutenberg.org/

as an *index* to get a corresponding sentence from generated map. And then, instead of regular primes, sequence of sentences is sent over the network. On the Bob's side reverse algorithm is used on the received sequence of sentences and the actual values of sending primes are shown to Bob. Bob calculates *b*, encrypts it the same way as Alice did with *a* and *p*, and send it to her. Also, he generates his key using the *b* value. Alice receives *b*, decrypts it and calculates the key. At the end, both sides have the same key - $b^a \bmod p = a^b \bmod p$.

To simulate the real conversation, sentences have to be sent with some delay. In the testing phase, this delay is set at 3 seconds. The plan is to do experimental research to determine the dependence of the time needed to type a sentence and the user's speed of typing. These results will be used to make this automated conversation more similar to the real conversation as it possibly is. In the current version (with Skype API), to complete the key exchange process, it takes around 10 minutes.

The result of successful key exchange between Alice and Bob is shown in Figure 3.

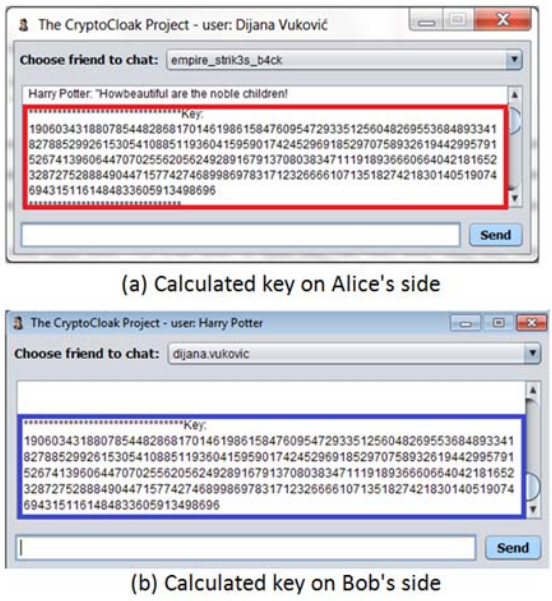

Figure 3. The CryptoCloak application

IV. RELATED WORK

The act of concealing data in plain sight is known as *steganography* [8]. Using *steganography* in network provides a possibility to carry on hidden information over the Internet as innocent Internet traffic. Authors in [8] presented results of their research in the field of network steganography to show how the network steganography can be used to exploit a common use of the Internet. These results can be helpful against violation of privacy. First presented research points out how the fact that Skype sends silence packets during the voice communication, when two sides have conversation break, can be exploited to send hidden information. The ***SkyDe*** program was developed to illustrate this. In their second research, the weakness of *BitTorrent*, that *BitTorrent* user often shares a data file (or pieces of the file) with many recipients at once, were used for encoding classified information in *BitTorrent* transactions. The ***StegTorrent*** program was developed to illustrate this. Third presented research results with the ***StegSuggest*** steganography program targets the feature Google Suggest, which lists the 10 most popular search phrases given a string of letters the user has entered in Googles search box. Using this fact secret information between two sides in communication will be exchange using Google Suggest. The forth presented research is related to Wi-Fi Networks weaknesses, on networks that use the data-encoding technique known as orthogonal frequency-division multiplexing (OFDM). They called this steganography method *Wireless Padding*, or ***WiPad***.

Internet censorship by government becomes an increasingly common practice worldwide. Between Internet users and censors now the arms race is started. For encrypted conversation over the Internet a lot of applications can be found. The most known one is **Tor** [12]. At the beginning Tor was a military project. It can be described as a "network of virtual tunnels". Tor provides protection from a common form of Internet surveillance known as "traffic analysis" by distributing transactions over several places on the Internet. This idea is similar to twisty - hard to follow route in order to throw off somebody who is tailing you. As a camouflage proxy for Tor, **StegoTorus** was developed. StegoTorus improves the resilience of Tor to fingerprinting attacks and delivers usable performance [13].

Cryptocat (open source software) [14] uses modern web technologies to provide easy to use, accessible encrypted chat. It is developed as plug-in for most popular web browsers. Chat conversation is encrypted before sending — even the Cryptocat network itself can't read it.

CryptoCloak is not a steganography. It does not embed the information in any other existing information. It produces a fake real-time, dynamic cheap chat and there it embeds the secret information. Messages sent via CryptoCloak application are not encrypted, as it is the case for Cryptocat, and they will not be detected by spying engines as suspicious (Figure 1).

V. FURTHER WORK

At the beginning of 2014, Skype API is retired and CryptoCloak needs to change the instant messenger used for chat communication over it. At the moment, chat communication is not working well, but there are plans to fix it in the future by providing a new API. *Facebook chat* use XMPP protocol for communication and the plan is to make a new version of the CryptoCloak using the XMPP protocol for sending chat messages instead of using Skype. The Smack library will be used [15].

In the next phase of implementation AES for encrypting/decrypting communication between Bob and Alice after the successful key exchange will be developed. To make conversation more look as a real chat between two persons, without sending sentences from the book, we have to find something like chat log archive or create one. Possibility that clients have simple databases of common sentences used for cheap chat communication instead of loading text files over the Internet will be added. In some further versions of application the plan is to add a possibility of choosing a topic for a chat.

To avoid dependency of other instant messengers API's, there is a possibility of implementing new instant messenger as a Java applet client.

VI. CONCLUSION

Privacy of individuals should not be threatened in any case, especially not in communication over the Internet. The CryptoCloak project has the aim – protection against surveillance in the chat communication. Diffie Hellman key exchange over the network, without sending a sequence of bytes, will not be detected by traffic analysis tools. The CryptoCloak is now in the testing phase, and some additional features may be added in the future.

ACKNOWLEDGMENT

I would like to express my very great appreciation to my supervisor, Professor Danilo Gligoroski, idea creator of the CryptoCloak project, for the given opportunity to be a part of it. Also, advice given by my co-supervisor Zoran Djuric has been a great help in implementation of the CryptoCloak application.

REFERENCES

[1] (Online resource) "NSA collecting phone records of millions of Verizon customers daily" (Available on: http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order).

[2] (Online resource) "UK gathering secret intelligence via covert NSA operation" (Available on: http://www.theguardian.com/technology/2013/jun/07/uk-gathering-secret-intelligence-nsa-prism).

[3] (Online resource) "Edward Snowden and the NSA files – timeline" (Available on: http://www.theguardian.com/world/2013/jun/23/edward-snowden-nsa-files-timeline).

[4] (Online resource) "The US government has betrayed the internet. We need to take it back." (Available on: http://www.theguardian.com/commentisfree/2013/sep/05/government-betrayed-internet-nsa-spying).

[5] (Online resource) Skype API For Java (English) (Available on: http://skype.sourceforge.jp/index.php?Skype%20API%20For%20Java%20(English)).

[6] (Online resource) "Cryptographers Have an Ethics Problem" (Available on: http://www.technologyreview.com/view/519281/cryptographers-have-an-ethics-problem/?goback=%2Egde_88033_member_273862254#%21).

[7] (Online resource) "ACM Code of Ethics and Professional Conduct" (Available on: http://www.acm.org/about/code-of-ethics).

[8] (Online resource) "4 New Ways to Smuggle Messages Across the Internet" (Available on: http://spectrum.ieee.org/telecom/security/4-new-ways-to-smuggle-messages-across-the-internet).

[9] (Online resource) " Skype says my application will stop working with Skype in December 2013, why is that?" (Available on: https://support.skype.com/en/faq/FA12349/skype-says-my-application-will-stop-working-with-skype-in-december-2013-why-is-that).

[10] Brian Martin, "Opposing Surveillance", IEEE Technology and Society Magazine, 29 (2), pp. 26-32, summer 2010.

[11] Travis D. Breaux and Catherine B. Lotrionte, "Towards a Privacy Management Framework for Distributed Cybersecurity in the New Data Ecology", in Proceedings of Conference Technologies for Homeland Security (HST), pp. 6-12, Waltham, MA, 2011.

[12] (Online resource) "Tor Project" (Available on: https://www.torproject.org/).

[13] Zachary Weinberg et.al, "StegoTorus: a camouflage proxy for the Tor anonymity system", in Proceedings of ACM CCS '12, pp. 109-120, 2012.

[14] (Online resource) "CryptoCat" (Available on: https://crypto.cat/).

[15] (Online resource) "Smack API" (Available on: http://www.igniterealtime.org/projects/smack/index.jsp).