Steganographic Method Based on Keyword Shift

Yong WANG, Qichang HE, Huadeng WANG, Bo YIN
School of Computer and Control
Guilin University of Electronic Technology
Guilin 541004, China
E-mail: hellowy@126.com

Shaoling DING
Institute of Information Technology
Guilin University of Electronic Technology
Guilin 541004, China
E-mail: dingshaoling2002@126.com

***Abstract*—By borrowing ideas from a cryptographic algorithm of low key authentic degree, a novel steganographic method based on keyword shift is presented. The master key of the method is to shift the sensitive keywords in the text. The conditions to guarantee the reversibility of the method are analyzed and found out, the serviceability of the method in some situations is pointed out. The strong points and weak points of the method are analyzed.**

Keywords- steganography; rubber-hose; digital signature; subliminal channel; information hiding

I. INTRODUCTION

Steganography is a technology and science about information hiding, which can ensures any unauthorized receivers can not discover the secret message except the authorized receiver. Nowadays information hiding technology mainly depends on large capacitance files such as images, audio files and videos files [1]. Another special information hiding technology is subliminal channel proposed by Simmons in 1983, in a narrow sense the subliminal channel is using digital signature to realize the information hiding [2], nevertheless subliminal channel can be sealed, and meanwhile the subliminal information using the digital signature is often very short. It is obvious that information hiding either depends on large files as carrier or merely transfers short messages. Covertext is very larger than stegotext, the utilize efficiency is low. Once large files and digital signature are forbidden to be sent, the information hiding of the secret can not be realized whereas information hiding and subliminal channel problems are always aiming at the prisoner problem, supervisor can forbid any transmission that may hide secret information absolutely. This paper proposes a new kind of steganography by borrowing ideas from a cryptographic algorithm of low key authentic degree.

II. PRINCIPLE OF THE KEYWORDS SHIFT STEGANOGRAPHY

We have designed a cryptographic algorithm against rubber-hose attack which adopted a method similar with doing multiple-choices questions [3]: this method has a database of keywords, every keyword is grouped with the keywords which are homoionyms or antonyms of this keyword , for instance, sunny is grouped with rainy and snowy, today is grouped with tomorrow and yesterday. Just like doing the multiple-choices questions, when encrypting, the keywords are replaced by an extend item, for example, keyword ‘rainy’ is replaced by an extend item ‘[(a)sunny (b)rainy (c) snowy]’, here symbol ‘[’ and ‘] ’ express the beginning and the end of an extend item, symbol ‘(’ and ‘)’ express the beginning and the end of a number of every keyword. In real encryption, these symbols should not appear in the text of plaintext files so that decryption is feasible and exclusive. The numbers used to sign keywords are consecutive integers from 0 to n-1 like “abcd” in the multiple-choices questions which can ensure the information is secretive and we can get different plaintexts when using different keys to decrypt, here n is the number of keywords in one group.

The character of this algorithm is that pseudokeys can be easily found, and the meaning of decrypted plaintexts using pseudokeys may be similar to or opposite to that of the right plaintext, thus it is easy to mislead the attacker. While traditional cryptographical algorithm is hard to find pseudokeys, when it comes to rubber-hose attack, there is defect in traditional cryptographical algorithm. If the key holder gives a key using which attacker can decrypt and get a meaningful text, then the attacker may believe that the key is right. This traditional algorithm has high authentic degree, contrarily the previous algorithm can easily find pseudokeys, so the authentic degree of the key of the algorithm is low [4].

We use similar method to hide information. Unlike encryption, steganography should be disguised as ordinary unencrypted communication, so the extend item cannot appear in the covertext and the covertext should like normal text. Therefore in steganographic method the sensitive keywords should be directly replaced by other keywords. When hiding, sensitive keywords are identified, according to key that receiver and sender shared and the initial number of the keyword in database we can compute another covertext number, then this number is used to ascertain which keyword should replace the sensitive keyword correspondingly. In the example above, “today is sunny” may be shifted to “tomorrow is sunny”, which has a misapprehend meaning.

III. DESIGN OF THE STEGANOGRAPHIC METHOD BASED ON KEYWORDS SHIFT

For this method shift the keyword of the text content of the stegotext file but not the file, then the step should include opening the document(or file), for example opening a word or text document, reading its text content of the document, then shift the keyword, saving the file at last.

In this paper we do not focus on the opening, reading and saving the file. We just consider the part of shifting keyword

978-1-4244-5265-1/10/$26.00 ©2010 IEEE

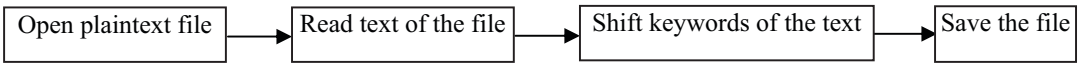

Figure 1. Flow Chart of Steganographic Method

of the text content. The processes of this part of the steganographic method are listed as follows:

The steganographic system scans the text read from the file and finds out the keyword in the text one by one according to the keywords database. If a word is a keyword in the keywords database then system reads the number 'n' of the corresponding group of the keyword in the database which is the number of the keywords of the group and 'a' which is the serial number of this keyword in the corresponding group in the database. The keyword will be replaced by another keyword in the group. Which keyword in this group will replace the original keyword is determined by the secret key. If the receiver and sender have shared a secret key k, we can use stream cipher algorithm to determine how to replace the keywords. If the largest number of keywords is not larger than 2^m, then we orderly choose m bit digits in stream cipher sequence which is generated by the stream cipher algorithm and the secret key when finding a keyword. Suppose at the s time, the value we transfer the corresponding m bits of the stream cipher sequence to decimal digits is d_s, the number of the keyword in the corresponding group which replaces No. s original keyword is gained by the following formula:

$b=(a+d_s) \bmod n$.

In this way steganographic system finds the keywords which replace original keywords in stegotext and generates the covertext.

The extraction of stegotext is similar, the words are the same with covertext if they are not keywords, the keywords are replaced according to the secret key orderly. This process is opposite to the hiding process.

Because of the same key, the receiver and sender will get the same stream cipher sequence if they adopt same stream cipher algorithm, for code sequence value correspondingly in the s time d_s , after querying keywords database we can find number b, then using $a=(b-d_s) \bmod n$, after querying keywords database about number a, we will gain keyword correspondingly and extract stegotext.

In order to be synchronous, either in steganographic process or in extraction we compute and shift keywords according to the text orderly. The purpose of adopting stream cipher algorithm is to avoid attacks using stegotext and covertext to find keys. Although adopting modular arithmetic in number computation process, stream cipher algorithm can efficiently prevent potential attack [5].

IV. Reversibility conditions of the steganographic method

In the above steganographic method if a keywords is contained by another, for example, in keyword database there are keyword 'China' and 'People's Republic of China' or there is a multivocal keyword in two different groups, the reversibility is hard to guarantee because there maybe different extraction results. Considering these conditions we should improve the algorithm and restrict the keywords. In the algorithm we can stipulate that a keyword should appear in no more than one group and any keyword should not be contained by another.

Steganography is different from encryption in order to preserve no flaw of the disguised covertext. Ciphertext can be unmeaning, but covertext should have meaning. Therefore steganographic method should get rid of the symbols used in our cryptographic algorithm that can avoid different interpretations when decrypting for the symbols differentiate the shifted words and unshifted words availably. As steganographic method gets rid of the symbols, that cause the following problems: the shifted words and unshifted words in the covertext may make up a new keyword that may generate a stegotext different from the original stegotext when extracting from the Covertext, although this case rarely happens, the deep reason of this problem is intersection of keywords. If there is no intersection between keywords, this problem will be solved. To this problem there are some solutions: 1). As this case rarely happens so we do not need to give limitations about keywords that require there is no intersection between two keywords in the database. We send messages if extracted stegotext is the same with the original stegotext. If not, stegotext should be improved or abandoned. 2). Searching keywords in series to find if there are two or more result, for example if there is " is not her" in the covertext, if both "is not" and "not her" are keywords in the database, then system gives two kinds of extract results. 3). Restrict keywords to avoid intersection. This solution is more efficient to English words, whereas in Chinese this limitation may reduce the number of keywords greatly and effectiveness of the method. There are other methods by coding to ensure the reversibility of the method.

According to the method, in the covertext there are no complete keywords in unshifted paragraph, limitations we give above can make sure there are no intersections between any keywords. By the above methods, we efficiently avoid or solve the keywords intersection between unshifted paragraph and shifted part, so there are no misjudgments in other keywords and it can ensure the reversibility of our steganographic method.

V. Applications of the steganographic method

The steganographic method can be applied as subliminal channel, when two prisoners try to communicate with each other and do not want guards who can see the letters know the true meaning of the letter, then two prisoners could adopt steganographic method to shift keywords in their stegotext, then turnkey can only see the covertext whose meaning may be opposite or similar to the stegotext. That can mislead the turnkey.

This steganographic method is more suitable to encrypt communications or falsify communications between computers or systems. For example, in internet, according to network protocols and communication mechanisms, we can

shift keywords such as URL, IP, commands, file names etc. The covertext also matches the mechanisms and protocols accordingly, so it can be processed availably and seems to be right. There are two uses for this: Firstly, this method can mislead the sniffer. Secondly, an attacker on the internet may falsify data using this steganographic method, and receiver is hard to find the data is changed and that lead to false operation of the computer or system.

VI. CONCLUSIONS

A steganographic algorithm based on keywords shift is proposed, which is quite other than the conventional steganographic method. Compared with conventional steganographic method, this steganographic method has its strong and weak points. This method has a value of a wide range of applications. By using similar method we can design many relevant steganographic and cryptographic algorithms. This algorithm can be used in steganographic method and subliminal channel, also can be used as cryptographic method.

ACKNOWLEDGMENT

This research was supported by Science and Research Foundation of Guangxi Ministry of Education (200911MS88), Guangxi Natural Science Foundation (0640171) and National Undergraduates Innovating Experimentation Project.

REFERENCES

[1] I.J. Cox, M.L. Miller, J.A. Bloom, J. Fridrich and T. Kalker, "Digital Watermarking and Steganography" (Second Edition), Morgan Kaufmann, 2008

[2] G J Simmons; The prisoner' problem and subliminal channel, Advance in Cryptology: Proceedings of Crypto'83, New York, 1984, pp. 51-67.

[3] Yong WANG, Security of One-time System and New Secure System [J],Netinfo Security, 2004, 7, pp. 41-43

[4] Yong Wang, Study of Some Problems of Quantum Cryptography and Theoretical Security of Cryptosystem [D], Southwest Jiaotong University, 2005

[5] Bruce Schneier , Applied Cryptography Second Edition: protocols, algorithms, and source code in C, John Wiley & Sons, Inc, 1996