

**Military Technical College  
Kobry El-Kobbah,  
Cairo, Egypt**



**8<sup>th</sup> International Conference  
on Electrical Engineering  
ICEENG 2012**

## **New Steganographic Method for Marking IP Stego Datagrams Based on the IP ID Field**

*By*

Manal A. Shehab

Noha O. Korany\*

### **Abstract:**

Stego datagrams could be transmitted as a continuous traffic stream or they could be marked and mixed with normal datagrams, then transmitted from the steganogram sender to the steganogram receiver. The steganogram receiver would check for the mark to identify the stego datagrams and extract the hidden data from them.

This paper suggests a new steganographic marking method with two scenarios to mark the IP stego datagrams using the IP ID field. This method could be used to mark the IP datagrams only and it requires from the packet source to be the steganogram sender. If this method is used to mark the IPv6 packets, then the IP packet fragmentation is required by the packet source.

### **Keywords:**

Stego, steganographic, data hiding, marking and IP ID field

\* Electrical Engineering Department, Faculty of Engineering, Alexandria University - Egypt

## **1. Introduction:**

There are two ways to transmit the stego packets from the steganogram sender (SS) to the steganogram receiver (SR). Stego packets could be transmitted as a continuous packets stream, or they could be marked and mixed with normal packets so that the SR could check for the mark to identify the stego packets and extract the hidden data from them. Transmitting a continuous stream of stego packets is easier in checking the hidden data than transmitting a mixed stream. The first case is easier to implement and it has higher average rate for delivering the stego packets compared to the second case.

Stego packets which hide data in their IP headers (as in the IP ID field) could be transmitted as a continuous packets stream (i.e. without marking them) in the common covert channels or they optionally (but not preferred) could be marked and mixed with normal packets for transmission. Stego packets which hide data in their payloads should be marked and mixed with normal packets and they should be checked and extracted away from the received stream at the SR, because the payload is the legitimate carrier of the transmitted data.

There are two ways to extract the hidden data from the marked IP stego datagrams within a mixed stream. The first way is sniffing or taking a copy of the received traffic at SR by a traffic capture to analyze it and extract the hidden data without affecting the received traffic, this way is easy and recommended for stego datagrams which hide data in their headers. The second way is excluding the stego datagrams away from the received traffic, this way requires more processing, introduces abnormal behaviors at SR and could be noticed by a warden.

If the number of the packets which are required to carry the covert message is  $Z$  stego packets and these packets would be marked and mixed with  $D-Z$  normal packets (where  $D$  is the total number of transmitted packets in the traffic stream), then  $D$  must be greater than  $Z$ . Great is the number of the normal packets compared to  $Z$ , more difficult is for a warden to detect the  $Z$  stego packets, but more processing and resources usage are required to check for the  $Z$  marked packets within the  $D$  packets stream at the SR. This needs to be considered by the SS and the SR to determine the suitable range of the used  $(D-Z)$  normal packet.

Cauch et al. [1] used the IP ID and the fragment offset fields of the IPv4 header to carry the hidden data by a certain manner between intermediate nodes under the condition of not fragmenting the packets. He used the reserved flag bit of the stego packet for

marking. This marking method is easy to detect because the stego packet has 1 in their reserved flag which contradicts its normal value and can attract the warden detection.

In the fragment header extension of the IPv6 packets; the 8-bit reserved field and the 2-bit reserved field for future work could be used to carry hidden data, but this is not recommended because these fields are initialized to zero for transmission and ignored on the reception. So; any initial value other than zero for the reserved fields at the packet source presents abnormal behavior that could be detected by a traffic warden.

Mazurczyk [2] decided to embed the hidden data in the payloads of legitimate fragments and he proposed a procedure to mark the selected stego fragments to help the SR in distinguishing them. His marking procedure depends on inserting a pre-agreed identifying sequence (IS) in the payload of the stego fragment.

This paper explains the idea of Ashan that regards the network MTU to reengineer the packet size to use the DF bit to carry a single hidden bit [3], this strategy could be used for marking the IP stego packets. Then the paper suggests a new steganographic marking method that uses the IP ID field to mark the IP stego packets, the method is denoted by MA for simplicity [4]. The MA steganographic marking method has two proposed scenarios for marking the IP stego datagrams only as it uses the IP ID field of the IP header to carry the mark.

## **2. Marking IPv4 Stego Packets Using the Don't Fragment (DF) Bit:**

This marking method inherits its idea from Ashan [3] who suggested using the DF bit of the IP header by the SS to transmit one hidden bit to the SR. If both SS and SR are on the same network (as being in the same LAN) and they know the MTU of their network, they could re-engineer the packet to avoid its fragmentation so that the DF bit could be used to carry a single hidden bit. Due to the MTU consideration of this method; if both SS and SR are communicated through a large network as the internet, re-engineering the packet size with the MTU of the transmission links would be more complicated. This method couldn't be used with the IPv6 packets as both the IPv6 header and the IPv6 fragment header extension don't include the DF bit.

Considering the strategy of the IPv4 fragmentation process; the IP header could be suspicious or non-suspicious. Suspicious IP header can catch the attention of the network administrator due to including abnormal data when it is compared with normal ones, while non-suspicious appears as normal ones to deceive the network monitoring devices. Thus; non-suspicious IP header is appropriate for data hiding.

Table 1 shows an example of a suspicious IPv4 header (only interested fields are shown) as the packet has a small size and fragmentation is not allowed (since the DF bit is set), whereas tables 2.a and 2.b show examples of non suspicious IPv4 headers as they show packets with moderate size. Fragmentation is not allowed in table 2.a since the DF bit is set, while fragmentation is allowed in table 2.b since the DF bit is not set. Both packets in tables 2.a and 2.b are non-suspicious packets and they are appropriate to hide single bit regarding that the frequent changes in the DF bit presents an abnormal behavior that could attract the detection of the network traffic monitors. Thus; both SS and SR shouldn't make this communication so frequent [3].

Packet #	16-bit IP ID	3-bit Flags	13-bit Fragment Offset	16-bit Total Length Field
1	XXX..XX	010	000..0	21

**Table (1): Suspicious IPv4 header**

Packet #	16-bit IP ID	3-bit Flags	13-bit Fragment Offset	16-bit Total Length Field
2	XXX..XX	010	000..0	472

**Table (2.a): Non suspicious IPv4 header with DF = 1**

Packet #	16-bit IP ID	3-bit Flags	13-bit Fragment Offset	16-bit Total Length Field
3	XXX..XX	000	000..0	472

**Table (2.b): Non suspicious IPv4 header with DF = 0**

**3. New Suggested Steganographic Method (MA) for Marking the IP Stego Datagrams Based on Embedding a Mark in the IP ID Field:**

**3.1 The MA1 Scenario: Using a Selected Bit in the IP ID Field as a Marking Flag:**

In the MA1 marking scenario; one bit of the IP ID field is chosen and its value is pre-agreed by the SS and the SR to mark the stego datagrams so that the SR could identify them from normal ones in the received stream. The remaining bits of the IP ID field would be generated randomly to serve the uniqueness property of the IP ID field.

**3.2 The MA2 Scenario: Embedding a Mark Code in the IP ID Field:**

In the MA2 marking scenario; a pre-agreed mark code (MC) is embedded in the IP ID field of the IP stego datagram (and consequently its resulted fragments). Assume that Alice is the SS and Bob is the SR, they need to agree about the embedding direction and two keys which are the mark code 'MC' and the mark code reference pointer key 'k', where k is used to point to the first bit of the mark code within the IP ID field of the IP stego datagram. The following algorithm describes the generation and the embedding of the mark code in the IP ID field of the stego datagram:

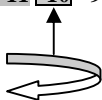
- 1- Assume that MC is the pre-agreed mark code and n is the minimum number of the bits which could represent the MC in its binary representation, where  $2 \leq n \leq 8$ . As more bits are in a frequently repeated mark code, as more attracting detection could be introduced.
- 2- Alice would convert the MC to its n-bits binary representation.
- 3- Alice would calculate the index x which points to the position of the first bit of the embedded mark code at the IP ID field.  
 For IPv4;  $x = k \pmod{16} + 1, 1 \leq x \leq 16$   
 For IPv6;  $x = k \pmod{32} + 1, 1 \leq x \leq 32$
- 4- Alice would embed the (n-bit) MC in the IP ID field starting from bit number 'x' in the clock wise or the anti- clock wise direction according to her previous agreement with Bob.
- 5- The remaining bits of the IP ID field would be generated randomly.

For example if Alice at the SS side agreed with Bob at the SR side to use the reference pointer key  $k = 26$  to embed a  $MC = 5$  in the clockwise direction within the IP ID of the IPv4 stego packets to mark them. Then the binary form of the used mark code is  $(MC)_2 = [101]$  and  $n = 3$ . Alice would calculate the index  $x$  from  $k$  as;

$$x = k \pmod{16} + 1 = 26 \pmod{16} + 1 = 11 \pmod{16}.$$

Assume that the 16-bit of the IP ID field are denoted as  $[i_{15}, \dots, i_1, i_0]$ , then Alice would embed the 3 bits of the MC  $[m_2, m_1, m_0]$  in the IP ID field starting from the bit number 11 which is  $i_{10}$  with its two followed bits in the clock wise direction and she would generate the remaining bits of the IP ID field randomly as follows:

The IP ID field  $[i_{15} i_{14} i_{13} i_{12} i_{11} i_{10} i_9 i_8 i_7 i_6 i_5 i_4 i_3 i_2 i_1 i_0]$  would be for the stego IP packet as:



[R R R **1 0 1** R R R R R R R R R R]

Where “R” is a random generated bit. Bob at the SR would consider all the received IP packets with the above IP ID form as stego packets.

To overcome the problem of detecting a long MC string which is frequently repeated, the following solutions are suggested:

- 1- Let the mark code MC consumes less bits [4 bits (i.e.  $n = 4$ ) for simplicity], and generate the remaining bits of the IP ID randomly to increase the randomization and the uniqueness of the IP ID field. If the length of the mark code is one bit (i.e.  $n = 1$ ), this returns the MA1 marking scenario in which the mark code acts as a flag bit in the IP ID.
- 2- Avoid sending stego packets consecutively over a detected time period. It's preferred to send one or some stego-packets, then send some normal packets which do not include the mark code.
- 3- Use a list of pre-agreed mark codes and switch between them to increase the difficulty of detecting them.

#### **4. Evaluation of the MA Steganographic Marking Method:**

Choosing the IP ID field to embed a mark utilizes the benefits of the IP ID nature as the IP ID field of a packet is the same as the IP ID field of its fragments, the IP ID field is unique and immutable in the packet journey between the source and the destination. Also the randomness nature of the IP ID value increases the difficulty for a warden to doubt if it carries a hidden data or a mark or not because the resulted IP stego datagrams would appear as normal ones. The MA steganographic marking method is resistant to detect by packet filtering and stateful inspection firewalls.

To apply the MA marking method with IPv6 stego datagrams, IPv6 packet fragmentation is required by the packet source to include the IPv6 fragment header extension which contains the IP ID field. The MA1 steganographic marking scenario provides more uniqueness for the IP ID field than the MA2 scenario. MA2 is easier to detect than MA1 especially if the mark code lengths more bits or it is frequently repeated. The process of using certain mark code that consumes certain bits in the IP ID field to mark the IP stego-packets involves preventing the generation of this code in its bits' positions within the IP ID for the normal IP packets. If the length of the mark code is one bit (i.e.  $n=1$ ), it returns the MA1 marking scenario.

The MA steganographic marking method is not recommended to mark the packets which hide data in their IP ID field as it could be better to send them as a continuous stream. Even if it's decided to mix them with normal IP packets; using the MA method in this case would be better to mark the normal packets so that the SR could distinguish the stego unmarked packets. If the MA method is used to mark the packets which hide data in their IP ID field; it's necessary to avoid having common bits shared by both the used mark and the main hidden data within the IP ID field, and in this case; MA1 is preferred than MA2 and it's recommended to choose either the right or the left neighbor bit that follows the bits of the embedded hidden data to act as the marking flag because this facilities both the embedding and the extracting processes.

The MA steganographic marking method could also be used to mark stego datagrams which hide data in their payloads.

## **5. Conclusions:**

IP stego datagrams could be transmitted as a continuous data stream or they could be marked and mixed with normal datagrams. Stego packets which hide data in their payloads should be marked and mixing with normal packets at the SS side to be identified and excluded away by the SR side, because the payload is the legitimate carrier of the transmitted data.

The MA steganographic marking method is proposed and evaluated. Two scenarios are presented for MA to mark the IP stego datagrams within a mixed data stream. The MA1 marking scenario uses a pre-agreed bit within the IP ID field with a pre-agreed value as a marking flag, whereas the MA2 marking scenario embeds a pre-agreed mark code in a pre-agreed position within the bits of the IP ID field. To apply the MA marking method with the IPv6 stego datagrams, IPv6 packet fragmentation is required by the packet sender. The MA1 marking scenario is preferred more than the MA2 because it provides more uniqueness for the IP ID field and less detectability.

## **References:**

- [1] Enrique Cauich, Roberto Gómez , Ryouke Watanabe, "Data hiding in identification and offset IP fields", The 5th International School and Symposium of Advanced Distributed Systems (ISSADS), pp. 118–125, January 2005.  
[http://pdf.aminer.org/000/449/997/data\\_hiding\\_in\\_identification\\_and\\_offset\\_ip\\_fields.pdf](http://pdf.aminer.org/000/449/997/data_hiding_in_identification_and_offset_ip_fields.pdf)
- [2] Wojciech Mazurczyk, Krzysztof Szczypiorski, "Evaluation of Steganographic Methods for Oversized IP Packets", Telecommunication Systems: Modeling, Analysis, Design and Management, Volume 49: 3-4, ISSN: 1018-4864 (print version), ISSN: 1572-9451 (electronic version), Springer US, Journal no. 11235, March/April 2012.  
<http://www.springerlink.com/content/20xn5j76r25002t5/>
- [3] Kamran Ahsan, "Covert Channel Analysis and Data Hiding in TCP/IP", Master of Science, Department of Electrical and Computer Engineering, University of Toronto, 2002. <http://www.springerlink.com/content/20xn5j76r25002t5/>
- [4] Manal A. Shehab, "New Encryption and Steganographic Methods for Data Hiding in the IP Packets or Their Fragments", Master of Science in Electrical Engineering, Electrical Engineering Department, Faculty of Engineering, Alexandria University, Egypt, October 2011.

**Last access to the web sides was in 14/ 3/ 2012.**