Design of watermark trace-back system to supplement connection maintenance problem

Hwan-Kuk Kim*, Seung-Wan Han*, Dong-il Seo*, and Sang-Ho Lee**

* Department of Network Security Research, ETRI, Korea
(Tel : +82-42-860-3823; E-mail: {rinyfeel, hansw, bluesea}@etri.ac.kr)
** Department of Computer Science, ChungBuk National Uni, Cheong Ju, Korea
(Tel : +82-43-261-2253; E-mail: hlee@cbucc.ac.kr)

Abstract: Internet is deeply rooted in everyday life and many things are performed using internet in real-world, therefore internet users increased because of convenience. Also internet accident is on the increase rapidly. The security vendor developed security system to protect network and system from intruder. Many hackings can be prevented and detected by using these security solutions. However, the new hacking methods and tools that can detour or defeat these solutions have been emerging and even script kids using these methods and tools can easily hack the systems. In consequence, system has gone through various difficulties. So, Necessity of intruder trace-back technology is increased gradually. Trace-back technology is tracing back a malicious hacker to his real location. trace-back technology is largely divided into TCP connection trace-back and IP packet trace-back to trace spoofed IP of form denial-of-service attacks. TCP connection trace-back technology that autonomously traces back the real location of hacker who attacks system using stepping stone at real time.

In this paper, We will describe watermark trace-back system using TCP hijacking technique to supplement difficult problem of connection maintenance happened at watermark insertion. Through proposed result, we may search attacker's real location which attempt attack through multiple connection by real time.

Keywords: Trace-back, Hacking, Network Security,TCP Hijacking

1. INTRODUCTION

The flow of network communications is critical to operations in today's world. Businesses, governments, and educational institutions all depend on the uninterrupted movement of information. However, it is apparent from the increasingly common network security incidents that current network security approaches are not enough. The primary methods used today are reactive in nature, dealing with security threats only after they have affected networked computers. Perimeter-based defenses, such as firewalls and intrusion detection systems, require increasingly complex efforts to keep up with changes in technology, shifts in threat, and demands for new network services[1].

Network-based attacks have become a major concern to today's highly networked mission critical information system because today's Internet infrastructure is vulnerable to motivated and well-equipped attackers. Much work is being done to safeguard resources, detect an attack, and, if possible, attempt to thwart the attack. A more difficult problem is determining the origins of an attack. Accurate and reliable identification of attackers is currently extremely difficult because the network routing infrastructure is stateless and based largely on destination addresses. The attacker can generate offending IP packets masquerading as having originated almost anywhere, including from IP addresses that are not globally unique, such as those used to create private network. In general, attacks against sites or even the network infrastructure can be waged from the safety of complete anonymity[2].

So, Necessity of intruder trace-back technology is increased gradually. In recent years there has been renewal of interest in trace-back. This paper is intended as an investigation of use TCP hijacking technique for supplement connection maintenance problem happens to watermark trace-back system. The remainder of this paper is organized as follows.

Section 2 gives a brief overview of the trace-back technique. In section 3 and 4, we describe overview of watermark trace-back system and watermark trace-back system using TCP hijacking technique to supplement difficult problem of connection maintenance happens at watermark insertion. Finally, conclude in section 5.

2. TRACE-BACK TECHNOLOGY OVERVIEW

Trace-back technology is tracing back a malicious hacker to his real location. This is largely divided into TCP connection trace-back and IP packet trace-back to trace spoofed IP of form denial-of-service attacks. We are not concerned here with IP trace-back because connection maintenance problem happens in TCP connection trace-back.

2.1 TCP connection trace-back

A TCP connection trace-back is technique to trace origin location of hacker who attempts based on TCP connection by real time. Also, it is often called a connection chain trace-back.

The connection chain is defined as following.

Definition 1. connection chain
Given a series of computer hosts H_1, H_2, … , H_n (n>2), when a person(or program) sequentially connects from H_i into H_{i-1} (I=1,2..,n-1), we refer to the sequence of connections on <H_1,H_2,…,H_n> as a connection chain, or chained connection. The tracing problem of a connection chain is, given H_n of a connection chain, to identify H_{n-1}, … H_1 [3]

Here, connection chain means that TCP connection C_1 is created between two systems - H_0 and H_1 , if computer H_0 logs in to different system H_1 through network in Fig.1.

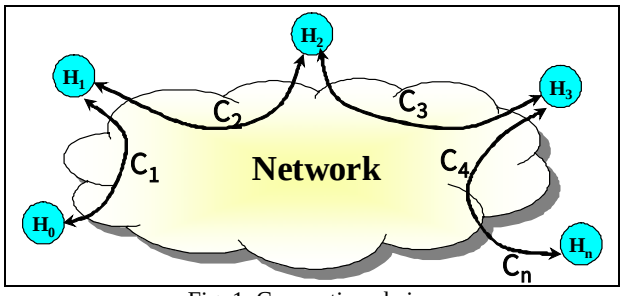

Fig. 1. Connection chain

A TCP connection trace-back technology is categorized by a host base connection trace-back and a network base connection trace-back.

A host base connection trace-back is a technology to trace with various information of log happen in host installing module for trace-back to hosts on the Internet. However, trace-back module should be placed to all hosts on the internet to achieve trace-back, and have shortcoming that trace-back is impossible if trace-back can not get information by certain problem occurs in system in path. To apply to current internet environment is almost impossible because of such problems.

A network base connection trace-back technology is placed in position that trace-back module can confirm packet pass through network and extract information that can achieve trace-back from packets. Now, most of proposed methods are adopting method that achieve trace-back extracting connection of connection chain correspond to attack connection at position can confirm packet pass through network. However, trace-back system was not proposed applying a network base connection trace-back technology to the current internet environment.

Only, it is circumstance that only algorithm is brought that can judge that belong to connection such as attack connection though some information is tacken advantage of from packets through network. While this shares various connection information that get from packets that pass network with trace-back systems that exist to network, order relation and synchronization of created information are very difficult because problem must possess information about all connections that occur in network continuously can happen[3].

2.2 IP trace-back

IP trace-back is to identify the true IP address of a host originating attack packets. Existing IP trace-back methods can be categorized as proactive or reactive tracing.

Proactive tracing prepares information for tracing when packets are in transit. If packet tracing is required, the attack victim(or target) can refer to this information to identify the attack source. Two proactive tracing methods – packet marking and messaging – have been proposed.

Reactive tracing starts tracing after an attack is detected. Most of the methods trace the attack path from the target back to its origin. The challenges are to develop effective trace-back algorithms and packet-matching techniques. Various proposals attempt to solve these problems. – hop-by-hop tracing, hop-by-hop tracing with an overlay network, IPsec authentication.[4].

3. PACKET WATERMARK TRACE-BACK

We shall now look more carefully into packet watermark trace-back. There are various kinds at TCP connection trace-back, one is packet watermark trace-back. Packet watermark trace-back is element technology to do trace-back about attack of purpose that hacker do not inform own system IP address via several systems of connection chain form.

3.1 Considerations

In conceptual, packet watermark is could use small information to identify one connection uniquely. Packet watermark must easily insert and extract hided to network application attacker. Therefore, watermark belongs to the application layer and is application-specific.

The following is two considerations to archive trace-back using watermark.

First, use of data hiding technology so that packet watermark may not be seen to attacker. One challenge in generating watermark is how to make watermarks invisible to end-users. For text based network applications such as telnet and rlogin, this is in many ways similar to hiding data in text, which is much more difficult than hiding data in pictures or sounds.

Second, a watermark must be able to traverse multiple connections and remain invariant for correlation. In order to trace-back along the intrusion connection chain, a mechanism is needed to find and match adjacent connections that belong to the same connection chain. We refer to this adjacent connection matching mechanism as correlation[5].

3.2 Packet watermark trace-back system

For the moment let us look closely at packet watermark trace-back system. Fig.2.-(a), (b) is system configuration and operation flow of trace-back system using packet watermark.

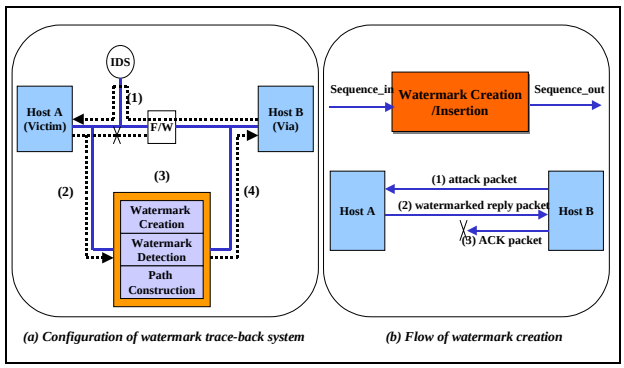

Fig. 2. System configuration and operation flow of watermark trace-back

A operation flow of watermark trace-back system is as following:

(1) detect attack packet (2) transmit reply packet (3) watermark creation and insertion (4) transmit watermark insertion packet (5) detect watermark packet (6) trace-back path construction

Fig. 3. shows process of TCP protocol's data transmission after connection synchronization through 3-way handshaking.

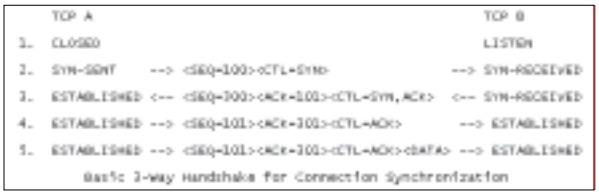

Fig. 3. TCP 3-way handshake

- Before watermark insertion :

$$Sequence_{(Host\ A)} = Sequence_{(Host\ A)} + Data\_Len \quad (1)$$

- After watermark insertion :

$$Sequence_{(Watermark)} = Sequence_{(Host\ A)} + Data\_Len + WM\_Len \quad (2)$$

- Result watermark insertion:

$$Sequence_{(Host\ A)} \neq Sequence_{(Host\ B)} <= Asynchronous \quad (3)$$

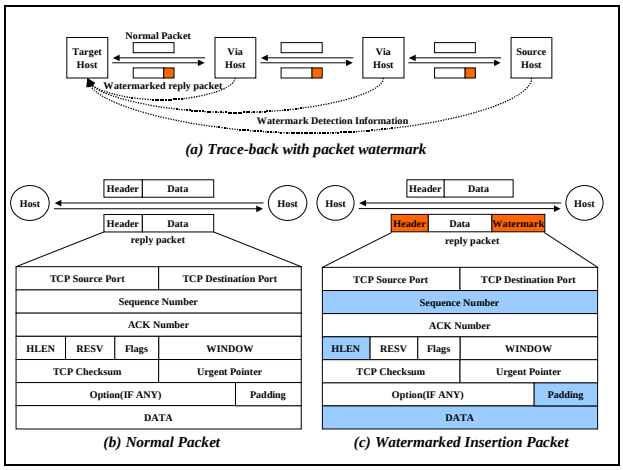

Fig. 4. Normal packet vs watermarked packet

When connection of both sides is established in TCP protocol and synchronous is achieved, data should be exchanged. However, packet watermark insertion does to change sequence number of reply packet in packet watermark trace-back system (Eqs. (1)~(3)). This becomes connection maintenance hardly because sequence number that is keeping from both end host gets into asynchronous state.

The problem then arises about such TCP protocol's sequence synchronous. This is the main problem of this paper. To supplement connection maintenance problem of both sides, we took advantage of TCP hijacking technology. TCP hijacking is techniques to make asynchronous state of SEQ, ACK number between server and client and get hijacked TCP session. This technology is active attack to use vulnerability of TCP protocol that can do redirection to flow TCP stream by own machine.

4. WATERMARK TRACE-BACK USING TCP HIJACKING

So, far we have seen watermark trace-back system and problems happens to asynchronous. We shall discuss trace-back system using TCP hijacking to supplement connection asynchronous problem.

The watermark trace-back system using TCP hijacking consists of connection hijacking block, watermark creation block, watermark detection block and path construction block. Here we limit the discussion to connection hijacking block within the scope of this paper.

The following Fig 5. is system configuration and operation flow .

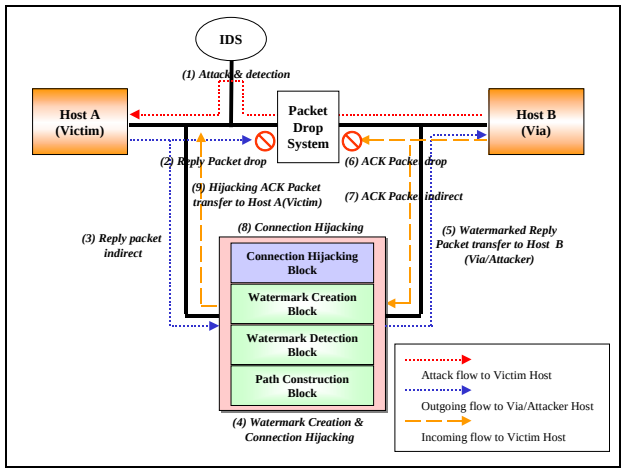

Fig. 5. System configuration and operation flow of watermark trace-back system using TCP hijacking

(1) If attack packet that happen in attacker host B are detected in host A
(2) drop reply packet that go out from host A to host B
(3) bypass reply packet to trace-back system
(4) create/insert watermark on reply packet and stored connection information about detected attack on connection hijacking block
(5) then transmit watermarked packet to via/attacker host B.
(6) drop ACK packet incoming after watermarked packet transmit.
(7) bypass ACK packet
(8) compare bypass ACK packet and attack information stored in connection hijacking block, and then process connection hijacking related attack.
(9) transmit hijacked ACK packet to Host A(Victim)

The connection hijacking block manages connection ID, source/destination address, watermark packet size information and process incoming/outgoing connection sequence related attack connection.

In Fig.6., the *INseq, OUTseq* is sequence number before/after watermark insertion of connection related attack. Fig 6.-(a) shows operation of watermark insertion without connection hijacking. Fig. 6.-(b) shows operation of proposed connection hijacking. The *New OUTseq* is reply packet sequence number outgoing to via /hacker host after watermark insertion. The value of *New OUTseq* is *INseq* added inserted watermark size at connection hijacking block.

$$New\ OUTseq = INseq + SIZE(WM) \quad (4)$$

In Eq(4), The *New INseq* is packets sequence number incoming to victim host after watermarked packet transmit. The *New INseq* is influenced sequence number by watermark insertion.

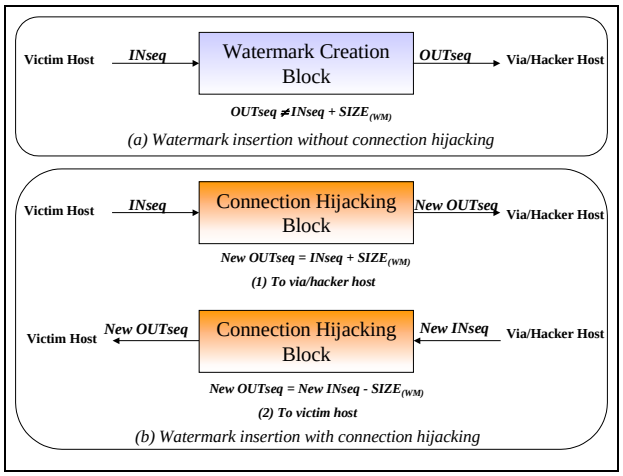

Fig 6. Connection hijacking block

When ACK packet come in connection hijacking block after watermarked packet transmit. In Eq(5), The value of *New OUTseq* is *New INseq* minus watermark size related attack connection :

$$New\ OUTseq = New\ INseq - SIZE(WM) \quad (5)$$

Therefore, such a solution of connection asynchronous problem happened at watermark insertion is linked with correlation between multiple connection chain previously section mentioned.

5. CONCLUSION

In this paper, we described watermark trace-back using TCP hijacking technique to supplement problem of connection maintenance happened at watermark insertion in watermark trace-back system. Through proposed result, we may search attacker's real location which attempts attack through multiple connection in real time.

Also, watermark trace-back system using TCP hijacking technique proposed in this paper can be utilized not only to seize a malicious hacker who compromises the system but also to deter curious script kids from trying to hack the system with curiosity.

However, there remains many problems which we must consider future work to apply to current internet environment: The problems utilize existent security tools(F/W, IDS), The problem of which trace-back system(at least, watermark detection module) has to be placed to all networks everywhere and secure exchange trace-back information between trace-back system and trace-back system on path had placed to many network.

ACKNOWLEDGMENTS

I wish to thank my colleagues for helpful comments on an draft of this paper.

REFERENCES

[1] T. J. Shimeall, C. J. Dunlevy, and Linda Pesante, “Challenges of Predictive Analysis for Networks” *International Journal of Control*, Vol. 23, No. 4, pp. 123-145, 1989.

[2] L. A. Sanchez, W. C. Milliken, A. C. Snoeren, F. Tchakountio, C. E. Jones, S. T. Kent, C. Partrige, and W. T. Strayer,“*Hardware Support for a Hash-Based IP Traceback*”, roceedings of the 2nd DARPA Information Survivability Conference and Exposition (DISCEX II), pp. 146-152, Anaheim, CA, June 2001.

[3] K. Yoda and H. Etoh, “Finding a Connection Chain for Tracing Intruders,” *6th European Symposium on Research in Computer Security*, ESORICS 2000 LNCS, France, 2000.

[4] T. Baba and S. Matsuda, “Tracing Network Attacks to Their Sources,” *IEEE Internet Computing*, pp. 20-26, March 2002.

[5] X. Wang, D. Reeves, S. F. Wu, and J. Yuill, “Sleepy Watermark Tracing: An Active Network-Based Intrusion Response Framework”, *Proc. of IFIP Conf. on Security*, Paris, 2001.

[6] Y.S.Choi, D.I.Seo, and S.W.Son, *“Trend of Traceback Technology :Focused to TCP Connection Traceback”*, ETRI A Weekly Technology Trend, Vol. 1078, pp. 13-25, 2003.