International Journal of Electrical, Electronics and Computer Engineering **3**(1): 173-176(2014)

ISSN No. (Online): 2277-2626

Effective Correlation in Stepping Stones by Flow Watermarking in Encrypted Attack Traffic

Vivek Patil and Ranjit Gawande

Department of Computer Engineering,
Matroshri College of Engineering and Research Center, Nashik,
University of Pune (MS)

(Corresponding author Vivek Patil)

(Received 05 May, 2014 Accepted 25 June, 2014)

ABSTRACT: In a network system many times it is seen that network intruders to hide their origin, they attacks through intermediate hosts called stepping stones. It is very important to observe and to correlate the flows of a stepping stone between the incoming and outgoing traffic, to identify the source of the attack behind the stepping stones. Sometimes it is possible that intruder encrypt the connection traffic to avoid attempts at correlation.

In this paper the correlation scheme based on watermarked signature is introduce. It is different from the previous timing-based correlation approaches. In the scheme, the traffic from victim back to the attacker of the bidirectional attack connections by slightly adjusting the timing of selected packets. By slightly adjusting the timing of selected packets this method embeds a unique watermark into the encrypted flows.

This method also overcome the most important drawback of host activity based methods is that the host activity collected from each stepping stone is generally not trustworthy. The attacker can easily modify or delete user login information. This makes host activity based correlation quite ineffective.

***Keywords*: Correlation, Perturbation, Robustness, Stepping Stones, Watermark, IPD.**

I. INTRODUCTION

In the working environment, network based attack is the major issue which is threat to the important information. And many times it is very difficult to identify the source of attack in a network. Since the attacker hides their identity and origin. As an example, they spoof the IP source address of the attack traffic. Since there are methods of tracing spoofed traffic, generally known as IP trace back [3], [4], [2], [10] have been developed to address this countermeasure.

Network-based intruders hide their identity by common and effective countermeasure way by connect through a sequence of intermediate hosts, it can be called as stepping stones, before attacking the final target. Let us consider an example, The hosts P,Q,and R are in network . An attacker at host P may Telnet or SSH into host Q, and from there the host P launches an attack on host R. That effect,s the attacking incoming packets of an connection are from P to Q are forwarded by Q, and become outgoing packets from Q to R. If the victim host R uses the method of IP traceback to determine the second flow originated from host Q, in that case the traceback will not be able to correlate that with the attack flow originating from host P. To trace attacks through a stepping stone, it is necessary to correlate the incoming traffic with the outgoing traffic at each stepping stone. This would allow the attack to be traced back to host P in above the example.

In earlier days the work on connection correlation was based only on tracking user's login activities at multiple hosts [12], [11]. In Later work the techniques of comparing the packet contents, or payloads, of the connections to be correlated were introduced [5], [6]. And most recent work has focused on the timing characteristics [8], [7], [1], [9] of connections, in to correlate encrypted connections.

Generally the timing based approaches are passive in nature because they observe and examine the network traffic. This paper introduces correlation of encrypted connections between intermediate hosts. The purpose of this method is to develop an efficient technique which is effective against random timing perturbation. And this correlation scheme is active since it embeds a unique watermark into the encrypted flows by slightly adjusting the timing of selected packets. The propose approach depend on watermarked based correlation is very much effective over the passive approaches. In this method the minimum packets are required to achieve better correlation in network.

The remainder of this paper is organized as follows. Section II Literature Review summarizes related work. Section III gives overview of Proposed System and Algorithm. Section IV describes Expected Results. Section V Describes the Conclusion Section.

II. LITERATURE REVIEW

The connection correlation approaches which are already existing are mainly focuses on host activity, connection content . Intrusion Detection System which works in distributed environment . The drawback of the host activity methods is that the host activity collected from each stepping stone is generally not trustworthy. Since the attacker is assumed to have full control over each stepping stone, and can easily modify user login information. This is not helpful to correlate based on host activity.

The different Content based correlation approaches like Thumb printing [5] which is short summary of content of a connection can be compared to determine whether two connections contain same text and therefore are likely to be part of same connection chain is content based correlation approach and SWT [6] is able to trace back to the trustworthy SWT guardian gateway that is closest to the source of intrusion chain, within single keystroke of the intruder. It require payload part of packets remains unchanged across stepping stones. However the attacker can easily transform the content by encryption at the application layer, these approaches are suitable only for unencrypted connections. To correlate encrypted traffic, in timing based approaches like ON/OFF-based [8], Deviation-based [7] and IPD-based [13]) . These methods only examines the arrival and departure times of packets, and use this information to correlate incoming and outgoing flows of a stepping stone. For in- stance, IPD-based correlation [13] has shown that 1) the important inter-packet timing properties of connections are preserved during transit across many routers and stepping stones and 2) the timing characteristics of interactive flows (e.g. telnet and SSH connections) are almost always unique enough to differentiate related flows from unrelated flows. The earlier timing based correlation approaches have proved to be effective in correlating encrypted connections. Donoho *et al.* [1] first investigated the theoretical limits on the attacker's ability to disguise his traffic through timing perturbation and bogus (padding, or chaff) packet injection.

III. PROPOSED SYSTEM AND ALGORITHM

A. Watermarking Model and Concept

Consider digital watermarking process [14], it involves the selection of a watermark carrier having the design of two processes embedding and decoding which are complementary to each other. First we collect the watermark signature, and then the watermark embedding process inserts the information by a slight modification of some property of the carrier. The decoding process of watermark extracts and decode watermark.

To analysis correlation between encrypted connections, use the inter-packet timing as the watermark carrier property of interest. The watermark embedded bit is guaranteed to be not corrupted by the timing perturbation. At the same time perturbation can be outside this range, the attacker may be altered embedded watermark bit .

Consider the network traffic where the number of packets flows in one direction, and packets are flows in intermediate hosts called steeping stones in a network, let consider arrival and departure time of each packet in steeping stone

Let consider, Tx and T1x be the arrival time and departure time respectively of a packet say Px of some stepping stones

And it is assumed that the normal processing and queueing time is required and it it considered as a delay which have to add in the stepping stone which is some constant say Const and is it greater than zero i.e. Const >0, and in that the attacker introduces extra delay say Dx to packet Px at the stepping stone, then we have

$$T1x = Tx + Const + Dx.$$

Now we calculate the two different inter packets delays

(i) Arrival inter packet delay

(ii) Departure inter packet delay

Arrival inter packet delay and Departure inter packet delay is calculated between the two packets say Px and Py

Arrival inter packet delay(AIPD) between Px and Py is as IPD (x,y)=Ty-Tx

Departure inter packet delay(DIPD) between Px and Py is as IPD1 (x,y)=T1y-T1x

AIPD or DIPD will be denoted as IPD when it is clear in the context. We further define the perturbation on IPD(x,y) by the attacker as the difference between IPD1(x,y) and IPD(x,y) : IPD1(x,y) – IPD(x,y) = Dy-Dx.

We use the timestamp of the xth and the yth packets to calculate IPD(x,y) or IPD1(x,y) even if there might be some packets reordered in the packet flow. Since we only use the timestamp of selected packets, the negative impact of using the "faulty" packet due to packet reorder is same to little random timing perturbation over the IPD.

If D is the delay that can be add by the attacker then the impact on IPD is Dy-Dx . If the delay is greater than zero then the perturbation is in between range -D to D is the perturbation range of the attacker. And method is most effective if we embed the watermark using inter packet delays from randomly selected packets [15].

IV. MODULES OF SYSTEM

A. Embedding and Decoding of Watermark bit

In registration process, we collect the watermark signature. watermarking involves the selection of a watermark carrier, and the design of two complementary processes: embedding and decoding. Embedding process involves insertion of the information by slightly modifying the property of the carrier and Decoding process involves detection and extraction of the watermark. Generally the delay between inter packet is the continuous value for that, we first quantize the inter packet delay before embedding the watermark bit.

For any non negative IPD we have to calculate the quantization of IPD with uniform quantization step size qs > 0 as the function

q(IPD , qs) = round(IPD,qs)

Let IPD is the Inter Packet Delay before watermark bit bw is embedded, and IPD1 denote the Inter Packet Delay after watermark bit bw is embedded. To embed bit w into an IPD, we slightly adjust that IPD such that the quantization of the adjusted IPD will have bw as the remainder when the modulus 2 is taken.

Given any IPD > 0; qs > 0 and binary digit bw, the watermark bit embedding is defined as function

e(IPD;bw; qs) = [q(IPD + qs=2; qs) + ¢] £ qs

where ¢ = (bw - (q(IPD + qs=2; qs) mod 2) + 2) mod 2.

The embedding of one watermark bit bw into scalar IPD is done through increasing the quantization of IPD+qs=2 by the normalized difference between bw and modulo 2 of the quantization of IPD+qs=2, so that the quantization of resulting IPD1 will have bw as the remainder when modulus 2 is taken. The reason to quantize IPD+qs=2 rather than ipd here is to make sure that the resulting e(IPD;bw; qs) is no less than IPD, i.e., packets cannot be output earlier than they arrive [15].

The watermark bit decoding function is defined as d(IPD1; qs) = q(IPD1; qs) mod 2

B. Correlation Analysis

To correlate encrypted connections, the propose to use the inter-packet timing as the watermark carrier property of interest. The embedded watermark bit is guaranteed to be not corrupted by the timing perturbation. If the perturbation is outside this range, the embedded watermark bit may be altered by the attacker. By embedding a unique watermark into the inter-packet timing, with sufficient redundancy, we can make the correlation of encrypted flows substantially more effective against random timing perturbations. One can also analyses the correlation of the watermark signatures and identify it's the positive or negative correlation, if positive occurs it detect it is the authenticated user otherwise, if negative occurs it detect it is an Intruder.

C. Watermark Tracing Model

The concept of watermark tracing is focuses on bidirectional traffic mainly from victim back to attacker . In this concept the bidirectional attacks are watermark with the backward traffic of bidirectional attack connection by adjusting the timing of selected packets . If the embedded watermark is both robust and unique, the watermarked back traffic can be effectively correlated and traced across stepping stones, from the victim all the way back to the attacker, assuming the attacker has not gained full control on the attack target, the attack Target will initiate the attack tracing after it has detected the attack. Specifically, the attack target will watermark the backward traffic of the attack connection, and inform across the network about the watermark. The stepping stone across the network will scan all traffic for the presence of the indicated watermark, and report [15].

To the target if any occurrences of the watermark are detected.

D. Parameter & Mapping Randomization

For the mapping we use technique of cryptography in which we use a secret key to generate a pseudo-random sequence of numerical values and add them in to the pixels in the watermarking area we can called it as a parameter randomization. And we can recover the original pixel values by the compound mappings and this technique is refer as mapping randomization. We may also combine this technique with the parameter randomization technique to enhance the security. This parameter exchange does not affect the effectiveness of lossless recoverability. Finally, the Authenticated user take the file in zip format with proper password.

V. EXPECTED RESULTS

1) By taking different timing perturbation and by using the non registered watermark signature in different timing perturbation the attacker can attack at each steeping stone by adding random maximum delay or by adding same delay to in each packet in stepping stones or by fixing some timespan between the packets at each stepping stone and check wether the intruder found or not.

2) By Embedding the non registered signature for embedding process by this we can identify the intruder.

A. Data set

We will be testing our system using following depend on following types of

The system is checking by flow of at least 400 packets .

DS-1 : The system is checking by flow of at least 400 packets

DS-2 : DS2 contains 1000 synthetic telnet

B. Result set

This scheme is quite effective and improved in the case of all the three uniformly distributed random perturbation, self-similar perturbation and batch releasing perturbation also and also in DS-1 and DS -2 .

VI. CONCLUSIONS

To trace out the origin of a attacker through stepping stones in a connected network is difficult problem This proposed technique is based on timing correlation which is useful against random timing perturbation. And due to unique watermark embedding into the inter packet timing we can make the system more effective .

REFERENCES

[1]. D. Donoho. et al. Multiscale Stepping Stone Detection: Detecting Pairs of Jittered Interactive Streams by Exploiting Maximum Tolerable Delay. *In Proceedings of the 5th International Symposium on Recent Advances in Intrusion Detection (RAID 2002):* LNCS-2516, pages 17–35. Springer, October 2002.

[2] M. T. Goodrich. Efficient packet marking for large-scale ip traceback. In Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002), pages 117–126. ACM, October 2002.

[3] S. Savage, D. Wetherall, A. Karlin, and T. Anderson. Practical Network Support for IP Traceback. In Proceedings of ACM SIGCOMM 2000, pages 295–306. ACM, September 2000.

[4] A. Snoeren, C. Patridge, et. al. Hash-based IP Traceback. In Proceedings of ACM SIGCOMM 2001, pages 3–14. ACM, September 2001.

[5] S. Staniford-Chen and L. Heberlein. Holding Intruders Accountable on the Internet. In Proceedings of the *1995 IEEE Symposium on Security and Privacy*, pages 39–49. *IEEE*, 1995.

[6] X. Wang, D. Reeves, S. F. Wu, and J. Yuill. Sleepy Watermark Tracing: An Active Network-Based Intrusion Response Framework. *In Proceedings of the 16th Internatinal Conference on Information Security (IFIP/Sec 2001),* pages 369–384. Kluwer Academic Publishers, June 2001.

[7] K. Yoda and H. Etoh. Finding a Connection Chain for Tracing Intruders. In Proceedings of the 6th European Symposium on Research in Computer Security (ESORICS 2000), LNCS-1895, pages 191–205. Springer-Verlag, October 2002.

[8] Y. Zhang and V. Paxson. Detecting Stepping Stones. In Proceedings of the 9th USENIX Security Symposium, pages 171–184. USENIX, 2000.

[9] A. Blum, D. Song, and S. Venkataraman. Detection of Interactive Stepping Stones: Algorithms and Confidence Bounds. *In Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID 2004).* Springer, October 2004.

[10] J. Li, M. Sung, J. Xu and L. Li. Large Scale IP Traceback in High- Speed Internet: Practical Techniques and Theoretical Foundation. *In Proceedings of the 2004 IEEE Symposium on Security and Privacy, IEEE, 2004.*

[11] S. Snapp. et al. DIDS (Distributed Intrusion Detection System) - Motivation, Architecture, and Early Prototype. *In Proceedings of the 14th National Computer Security Conference*, pages 167–176, 1991.

[12] H. Jung. et al. Caller Identification System in the Internet Environment. *In Proceedings of the 4th USENIX Security Symposium, USENIX, 1993*.

[13] X. Wang, D. Reeves, and S. F. Wu. Inter-packet Delay based Correlation for Tracing Encrypted Connections through Stepping Stones. In Proceedings of the 7th European Symposium on Research in Computer Security (ESORICS 2002), LNCS-2502, pages 244–263. Springer-Verlag, October 2002.

[14] I. Cox, M. Miller, and J. Bloom. Digital Watermarking. Morgan- Kaufmann Publishers, 2002.

[15]Xinyuan Wang, Douglas S. Reeves Robust Correlation of Encrypted Attack Traffic through Stepping Stones by Flow Watermarking, *IEEE Transactions on Dependable And Secure Computing*, Vol. **8**, No. 3, May-June 2011.