International Journal on Emerging Technologies **6**(2): 235-239(2015)

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

Error-Free correlation in Encrypted Attack Traffic by Watermarking flow through Stepping Stones

Vivek Patil
Department of Computer Engineering,
Matroshri College of Engineering and Research Center, Nashik, (MS), India

(Corresponding author: Vivek Patil)
(Received 29 October, 2015 Accepted 18 November 2015)
(Published by Research Trend, Website: www.researchtrend.net patilvivek4@gmail.com*)*

ABSTRACT: In network system it is very few times network intruders directly attack from their own system, many times intruders attacks through intermediate stepping stones. At that time it is important to check and correlate the flows of a incoming and outgoing traffic between stepping stones, to identify the source of the attack behind the stepping stones. Sometimes the attacker may encrypt the connection traffic to avoid attempts at correlation. To prevent attempts at correlation in correlating encrypted connections timing based correlation approaches have been shown to be quite effective. However, timing based correlation approaches are subject to timing perturbations and that may be purposefully introduced by the attacker at stepping stones. In this paper the watermark-based correlation scheme is introduce that is designed specifically to be robust against timing perturbations and it is different from the previous timing-based correlation approaches. In the watermark-based correlation scheme the backward traffic of the bidirectional attack connections by slightly adjusting the timing of selected packets. This watermark-based approach can be called "active" since it embeds a unique watermark into the encrypted flows by slightly adjusting the timing of selected packets and it does not make any limiting assumptions about the distribution or random process of the original inter-packet timing of the packet flow. This method also overcome the main and 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, delete or forge user login information as he has full control over each stepping stone, he/she can. This makes host activity based correlation quite ineffective. This proposed model does not require the timing perturbation to follow any specific distribution or random process and makes no assumption about the distribution of inter packet timing intervals. This model is provably effective against certain correlated random timing perturbation, and requires substantially fewer packets than passive approaches to achieve

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

I. INTRODUCTION

Generally the network based attacks are become very a serious threat to the sensitive information infrastructure on which we depend. And to avoid or stop network-based attacks, it is very critical to identify the source of the attack in a network. However attackers, hide their identities and origin, using a variety of countermeasures. 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 traceback [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 L, M, and N are in network . An attacker at host L may Telnet or SSH into host M, and from there the host L launches an attack on host N. That effects the attacking incoming packets of an connection are from L to M are forwarded by M, and become outgoing packets from M to N. If the victim host N uses the method of IP traceback to determine the second flow originated from host M, in that case the traceback will not be able to correlate that with the attack flow originating from host L. 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 correlation was based only on user's login activities tracking at multiple hosts [11-12].

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.

Timing based approaches are unassertive Since they examine the network traffic. This paper introduces mutual relation of encrypted connections between intermediate hosts. This method is useful to develop an efficient technique which is effective against random timing perturbation. And this method 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. 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

Existing approaches of correlation focuses on, content connection and host activity. Intrusion Detection System which works in distributed environment. The disadvantage of the host activity methods is that the activity of host collected at each stepping stone is not relaible. Intruder have control on every stepping stone, so that he/she can easily change user login information, which is not helpful to correlate based on host activity. Other 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. SYSTEM ARCHITECTURE AND ALGORITHM

A. Watermarking Model and Concept

In digital watermarking process [14], involves the selection of a watermark carrier which having the design of two different processes decoding and embedding which are corresponding to each other. This process is describe in following steps

1. Collect the exact watermark signature
2. The watermark embedding process inserts the information by modifying of some property of carrier.
3. Watermark extracts and decode watermark in decoding process.
4. 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.

In the network traffic packets are flows in between intermediate hosts called steeping stones in a network, Let , Tz and T1z be the arrival time and departure time respectively of a packet say Pz of some stepping stones And assumed that queueig and processing time is required and delay which need to add in every the stepping stone which is some constant say C and $C > 0$, and in that the attacker introduces extra delay say Dz to packet Pz at the stepping stone, then we have

$T1z = Tz + C + Dz.$

After that two different inter packets delays need to calculate first is Arrival inter packet delay and second is Departure inter packet delay .Arrival inter packet delay and Departure inter packet delay is calculated between the two packets say Pz and Py

Arrival inter packet delay between Pz and Py is as $IPD(z,y)=Ty-Tz$

Departure inter packet delay between Pz and Py is as $IPD1(z,y)=T1y-T1z$

The perturbation on IPD(z,y) by the attacker as the difference between IPD1(z,y) and IPD(z,y) : $IPD1(z,y) – IPD(z,y) = Dy- Dz.$

If there might be some packets reordered in the packet flow, the timestamp of the zth and the yth packets to calculate IPD(z,y) or IPD1(z,y) . 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-Dz.

If D>0 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]. Fig. 1 shows System Architecture.

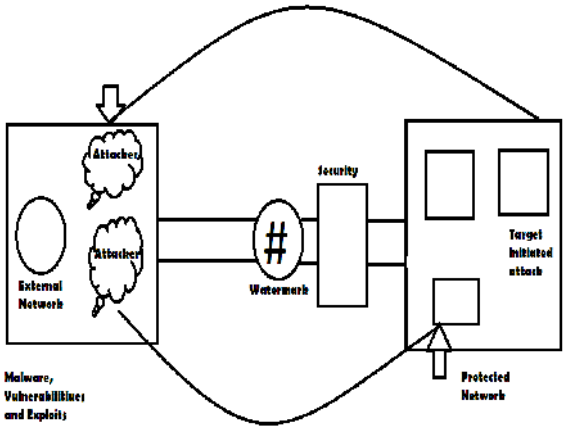

Fig. 1. System Architecture.

IV. MODULES OF SYSTEM

A. Watermark Bit Embedding and Decoding

In registration process, we collect the watermark signature. Watermaring process involves following steps

1.In embedding process information by modifiying of some property of the carrier

2. Decoding process extracts the watermark .

The delay between Inter Packet is the continuous value. We will first quantize the IPD before embedding the watermark bit. Given any IPD ip > 0, we define the quantization of ip with uniform quantization step of size q > 0 as the function

q(ip; q) = round(ip=q)

Let ip is the Inter Packet Delay before watermark bit bw is embedded, and ip1 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 ip > 0; q > 0 and binary digit bw, the watermark bit embedding is defined as function

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

where ¢ = (bw ¡ (q(ipd + q=2; q) mod 2) + 2) mod 2.

The embedding of one watermark bit bw into scalar ip is done through increasing the quantization of ip+q=2 by the normalized difference between bw and modulo 2 of the quantization of ip+q=2, so that the quantization of resulting ip1 will have bw as the remainder when modulus 2 is taken. The reason to quantize ip+q=2 rather than ipd here is to make sure that the resulting e(ip;bw; q) is no less than ip, i.e., packets cannot be output earlier than they arrive.

The watermark bit decoding function is defined as d(ip1; q) = q(ip1; q) mod 2

B. Correlation Analysis

By embedding a unique watermark into the inter-packet timing, with sufficient redundancy, we can make the correlation of encrypted flows substantially more robust 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 watermark tracing approach exploits the observation that interactive connections are bidirectional. The idea is to watermark the backward traffic of the bidirectional attack connections by slightly 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. 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. RESULTS

For checking the successful outcomes we can take following types of timing perturbations:

Random perturbation: In random timing perturbation the attacker at a stepping stone adds to each packet evenly distributed between 0 and the maximum delay, the daly can be chosen by the attacker i.e. random delay.

Self-similar perturbation: In self-similar perturbation, the attacker at a stepping stone adds to each packet random delay which is similar to itself.

Batch releasing perturbation : In the batch releasing perturbation the attacker at a stepping stone periodically buffers and holds all packets received within a certain time window, and once the time window has expired then forwards all the buffered packets at line speed.

A. Data set. We will be testing our system using following types of flows FS-1 : 121 SSH flows that have at least 600 packets and that are at least 300 seconds long,

FS1-Int : FS2 contains 1000 synthetic telnet

FS-2 : FS1-Int are interactive flows from type FS-1

The graph given below shows the comparison between packet splitting time for legal and illegal watermarks

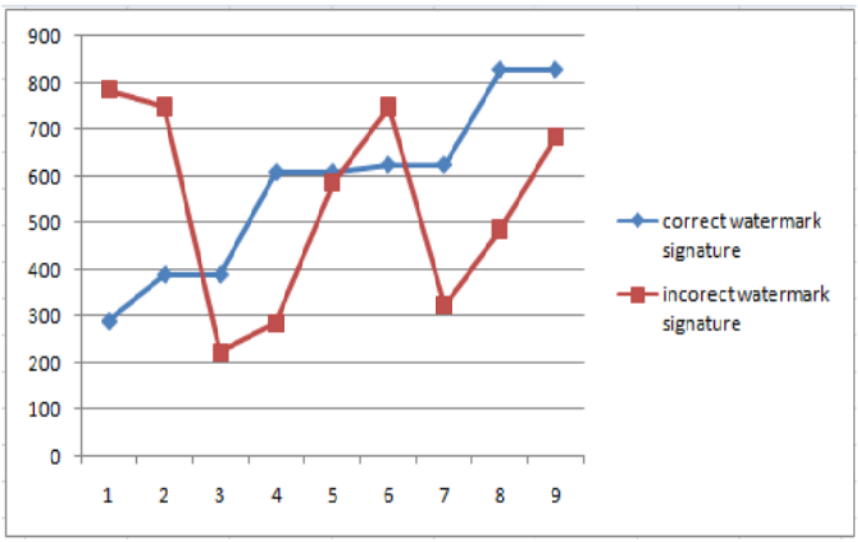

Fig. 2. Timing graph with respect to packets.

VI. CONCLUSION

Our active timing-based correlation approach is to deal with random timing perturbations. By embedding a unique watermark into number of splitted packet timing, with sufficient redundancy, we can make the correlation of encrypted more robust against timing perturbations. Our watermark-based correlation is effective against random timing perturbation. Our watermark-based correlation can achieve arbitrarily close to 100.

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.