An Effective Attack Analysis and Defense in Web Traffic Using Only Timing Information

THASNEEM M
II Year M.E(CSE)
Shree Venkateshwara Hi-Tech Engineering College, Gobi
tech.thasneem@gmail.com

S.RAMESH
Assistant professor(CSE)
Shree Venkateshwara Hi-Tech Engineering College, Gobi
judiciousramesh@gmail.com

Dr. T. SENTHIL PRAKASH
Professor & HOD
Shree Venkateshwara Hi-Tech Engineering College, Gobi
jtyesp@yahoo.co.in

Abstract – Attack against encrypted web traffic that makes use only of packet timing information on the uplink. This attack is therefore impervious to existing packet padding defenses. In addition, unlike existing approaches, this timing-only attack does not require the knowledge of the start/end of web fetches and so is effective against traffic streams. We demonstrate the effectiveness of the attack against both wired and wireless traffic, achieving mean success rates in excess of 90%. In addition to being of interest in its own right; this timing-only attack serves to highlight deficiencies in existing defenses and so to areas where it would be beneficial for virtual private network (VPN) designers to focus further attention. We propose to use the spatial correlation of received signal strength (RSS) inherited from wireless nodes to detect the spoofing attacks. We then formulate the problem of determining the number of attackers as a multiclass detection problem. Cluster-based mechanisms are developed to determine the number of attackers. When the training data are available, we explore using the Support Vector Machines (SVM) method to further improve the accuracy of determining the number of attackers. In addition, we developed an integrated detection and localization system that can localize the positions of multiple attackers. We evaluated our techniques through two test beds using both an 802.11 (Wi Fi) network and an 802.15.4 (Zig Bee) network in two real office buildings. Our experimental results show that our proposed methods can achieve over 90 percent Hit Rate and Precision when determining the number of attackers.

Keywords – Network privacy, timing-only attacks, traffic analysis, website fingerprinting

I. INTRODUCTION

In this paper we consider an attacker of the type illustrated in Figure 1. The attacker can detect the time when packets traverse the encrypted tunnel in the uplink direction, but has no other information about the clients’ activity. The attacker’s objective is to use this information to guess, with high probability of success, the web sites which the client visits. What is distinctive about the attack considered here is that the attacker relies solely on packet timestamp information whereas the previously reported attacks against encrypted web traffic have mainly made use of observations of packet size and/or packet count information.

Our interest in timing-only attacks is twofold. Firstly, packet padding is a relatively straightforward defense against attacks that rely primarily on packet size, and indeed is currently either already available or being implemented in a number of popular VPNs [2]. Secondly, alternative attacks based on packet counting [2], [3] are insensitive to packet padding defenses but require partitioning of a packet stream into individual web fetches in order for the number of packets associated with each web fetch to be determined, which may be highly challenging in practice on links where there are no clear pauses between web fetches. In contrast, packet timing-based attacks are not only largely unaffected by packet padding defenses but also, as we will show, do not require partitioning of the packet stream. Hence, they are potentially a practically important class of attack against current and future VPNs. While some work has been carried out using inter-arrival time information to classify the application (HTTP, IMAP etc.) [7], to our knowledge, there is no previous work reporting use of timing information alone to construct a successful attack against encrypted web traffic.

Fig. 1: Schematic illustrating attacker of the type considered. A client machine is connected to an external network via an encrypted tunnel (ssh, SSL, IPSec etc.). The attacker can detect the time when packets traverse the tunnel in the uplink direction, but has no other information about the client’s activity.

The main contributions of the present paper are as follows:

(i) we describe an attack against encrypted web traffic that uses packet timing information alone, (ii) we demonstrate that this attack is highly effective against both wired and wireless traffic, achieving mean success rates in excess of 90% over Ethernet and wireless tunnels and a success rate
of 58% against Tor traffic, (iii) we also demonstrate that the attack is effective against traffic streams i.e. back to back web page fetches where the packet boundaries between fetches are unknown.

In addition to being of interest in its own right, particularly in view of the powerful nature of the attack, this timing-only attack also serves to highlight deficiencies in existing defenses and so to areas where it would be beneficial for VPN designers to focus further attention. We note that, complementary to the present work, in [3] it is demonstrated that when the web fetch boundaries within a packet stream are known then an NGRAM approach using packet count together with uplink/downlink direction information is also sufficient to construct an effective attack against encrypted web traffic despite packet padding. Hence, we can conclude that (i) uplink/downlink packet ordering plus web fetch boundaries and (ii) uplink/downlink packet timing information are both sensitive quantities that ought to be protected by a secure encrypted tunnel. Packet padding does not protect these quantities. Directing defenses against these two sets of packet stream features therefore seems an important direction for future work.

II. EXISTING SYSTEM

An easy way to comply with the IJSRET journal paper formatting requirements is to use this document as a template and simply type your text into it. The attacker can detect the time when packets traverse the encrypted tunnel in the uplink direction, but has no other information about the clients’ activity. The attacker’s objective is to use this information to guess, with high probability of success, the web sites which the client visits. The attacker relies solely on packet timestamp information whereas the previously reported attacks against encrypted web traffic have mainly made use of observations of packet size and/or packet count information. Our interest in timing-only attacks is twofold. Packet padding is a relatively straightforward defense against attacks that rely primarily on packet size, and indeed is currently either already available or being implemented in a number of popular virtual private networks. Alternative attacks based on packet counting are insensitive to packet padding defenses but require partitioning of a packet stream into individual web fetches in order for the number of packets associated with each web fetch to be determined, which may be highly challenging in practice on links where there are no clear pauses between web fetches.

III. LIMITATIONS

The timing information about attack is only in the uplink direction. The attackers guess the information about which client visit the web sites. The attacker observes client port number and size. The attacker observes the direction and timing packet.

IV. PROJECT DESCRIPTION

When traffic is carried over an encrypted tunnel, such as a VPN, the packet source and destination addresses and ports and the packet payload are hidden. We also assume here that the tunnel pads the packets to be of equal size, so that packet size information is also concealed, and that the start and end of an individual web fetch may also be concealed e.g. when the web fetch is embedded in a larger traffic stream. An attacker sniffing traffic on the encrypted tunnel is therefore able only to observe the direction and timing of packets through the tunnel, i.e. to observe a sequence of pairs \( \{(t_k, d_k)\} \), \( k = 1, 2, \ldots \) where \( t_k \) is the time at which the \( k \)-th packet is observed and \( d_k \in \{-1, 1\} \) indicates whether the packet is travelling in the uplink or downlink direction. Since it will provide sufficient to mount an effective attack, we will assume a weaker attacker that can only observe the timestamps \( \{t_k\} \), \( k \in K_{\text{up}} := \{\kappa \in \{1, 2, \ldots \} : d_\kappa = -1\} \) associated with uplink traffic.

Fig. 2: Time traces of uplink traffic from 5 different Irish health-related web sites are shown. It can be seen that the web site time traces exhibit distinct patterns. The traces are shifted vertically to avoid overlap and facilitate comparison.

Figure 2 plots the timestamps \( \{t_k\} \) of the uplink packets sent during the course of fetching five different health-related web pages (see below for details of the measurement setup). The x-axis indicates the packet number \( k \) within the stream and the y-axis the corresponding timestamp \( t_k \) in seconds. It can be seen that these timestamp traces are distinctly different for each web site, and it is this observation that motivates interest in whether timing analysis may by itself (without additional information such as packet size, uplink/downlink packet ordering etc.) be sufficient to successfully de-anonymise encrypted web traffic. To gain insight into the differences
between the packet timestamp sequences in Figure 2 and, importantly, whether they are genuinely related to characteristics of each web page rather than to other factors, it is helpful to consider the process of fetching a web page in more detail. To fetch a web page the client browser starts by opening a TCP connection with the server indicated by the URL and issues an HTTP GET or POST request to which the server then replies. As the client parses the server response it issues additional GET/POST requests to fetch embedded objects (images, css, scripts etc.). These additional requests may be to different servers from the original request (e.g. when the object to be fetched is an advert or is hosted in a separate content-delivery network), in which case the client opens a TCP connection to each new server in order to issue the requests. Fetching of these objects may in turn trigger the fetching of further objects.

V. MODULE DESCRIPTION

The overall software contains are mainly four module
1. Node Creation
2. Gade
3. Silence
4. Idol

1. Node Creation
This Module Mainly to create a node with specific kind of information such as Node Name, IP Address, Port Number and those information’s are stored in the database. The wireless Communication established between the Nodes created on. Each and every node must be requesting to its server connected to the database. Multiple Adversaries may also obtain in the wireless communication.

2. Gade (Generalized attack Detection Model)
Generalized attack Detection Model is used to detect both the Spoofing attack and the number of attackers. Cluster based mechanisms are used to detect the Spoofing attack and the adversaries based on the RSS (Received signal strength). In GADE, the Partitioning Around Medoids (PAM) Cluster analysis method is used to perform attack detection. The problem of determining the number of attackers as a multiclass detection problem. Cluster-based methods to determine the number of attacker.

3. Silence
Silhouette Plot and System Evolution minimum distance of clusters, to improve the accuracy of determining the number of attackers. Additionally, when the training data are available, we propose to use the Support Vector Machines (SVM) method to further improve the accuracy of determining the number of attackers.

4. Idol
An integrated detection and localization system that can both detect attacks as well as find the positions of multiple adversaries even when the adversaries vary their transmission power levels. IDOL can achieve similar localization accuracy when localizing adversaries to that of under normal conditions. One key observation is that IDOL can handle attackers using different transmission power levels, thereby providing strong evidence of the effectiveness of localizing adversaries when there are multiple attackers in the network.

VI. CONCLUSION

To use received signal strength based spatial correlation, a physical property associated with each wireless device that is hard to falsify and not reliant on cryptography as the basis for detecting spoofing attacks in wireless networks. The spatial correlation of RSS inherited from wireless nodes for attack detection. The test statistic based on the cluster analysis of RSS readings. Our approach can both detect the presence of attacks as well as determine the number of adversaries, spoofing the same node identity, so that localize any number of attackers and eliminate them. Determining the number of adversaries is a particularly challenging problem. We developed SILENCE, a mechanism that employs the minimum distance testing in addition to cluster analysis to achieve better accuracy of determining the number of attackers than other such as Silhouette Plot and System Evolution that use cluster analysis alone. Additionally, when the training data are available, using Support Vector Machines-based mechanism to further improve the accuracy of determining the number of attackers present in the system.

ACKNOWLEDGMENT

I take great pleasure in expressing my profound gratitude to all those who have helped for the successful completion of the work.
To begin with, I offer my humble thanks at the sacred feet of the Almighty; the creator whose will to make this work a great success.
REFERENCES


AUTHOR PROFILE

Thasneem.M
ME CSE Student
Shree venkateshwara Hi-Tech
Engineering College, Gobi