An Effective and Feasible Traceback Scheme in Mobile Internet Environment

Shui Yu, Senior Member, IEEE, Keshav Sood, and Yong Xiang, Senior Member, IEEE

Abstract—Around one billion people access the Internet using their mobile phones today, and many of the mobile phones are prone to be compromised by hackers due to their inherent vulnerability. It is critical to identify these compromised mobile phones to effectively eliminate cyber attacks. However, we see few research works in the field. In order to address this desperate situation, we design a practical traceback framework to identify active compromised mobiles in the mobile Internet environment in this letter. In the proposed framework, we creatively use the IMEI number of mobile hardware as unique marks for the traceback purpose. Two-layer traceback tables are designed to collect global attack information and identify local attacking bots, respectively. Our analysis and simulation demonstrate that the proposed traceback method is effective and feasible, and it can identify every possible attacking mobile in the current mobile Internet environment with single packet marking.

Index Terms—Traceback, mobile Internet, attack source.

I. INTRODUCTION

It is pervasive to accessing the Internet through mobile phones nowadays. With the popularity of mobile phones and rich web based information, people heavily depend on mobile phones for Internet accessing, such as online news, email, online social networks, online music or TV, and so on. We usually name this as WAP (Wireless Application Protocol) applications. A WAP request needs to go through two networks, the cellular network and the IP network, to complete an Internet access. A bridge, WAP gateway, connects the two types of networks translating between the two sets of communication protocols. At the same time, mobile phones have become easy targets of cyber attackers due to their limited computing power and resources for protection. A large number of mobile phones are compromised by cyber attacker to commit malicious tasks. We usually call a compromised mobile as a bot or a mobile bot, and multiple bots controlled by the same hackers form a botnet or mobile botnet. Research has shown that mobile botnet develops dramatically in recent years [1].

It is hard to identify the sources of cyber attacks from mobile networks. Identifying attack sources is always a critical strategy to counter cyber attacks. This topic has been studied in the IP network since the 1990s. There are mainly three techniques for traceback: probability packet marking (PPM in short) [2], deterministic packet marking (DPM in short) [3], and entropy based traceback [4]. Most the previous PPM and entropy based methods share the same disadvantage of accuracy, while the DPM methods are hindered in practice by scalability. Recently, Yu et al. [5] proposed a marking on demand (MOD) traceback strategy based on the DPM mechanism in IP networks, which makes traceback practical and improves accuracy significantly. Moreover, a recent novel strategy of preventing DDoS attacks is also very interesting [6].

However, the traceback task in mobile Internet is seldom investigated. Bai [7] pointed out the difficulties of the traceback in the mobile Internet environment, and proposed to combine various logs to identify attack sources. Siddiqui et al. [8] touched this field through logging based traceback using bloom filters. However, these proposed methods are far from practical. Unfortunately, the MOD scheme (and most of the existing traceback schemes) cannot be applied in the WAP environment. In the existing traceback schemes, IP addresses are used as marks for traceback, therefore, we can maximum traceback to the WAP gateways, rather than specific mobile phones. Moreover, the mobility of mobile phones worsen the effectiveness of traceback.

In this letter, we propose a novel, effective and feasible method to traceback to individual mobiles who involve in attacks. Inspired by the marking on demand strategy, we designed a practical traceback scheme to identify active mobile bots with single IP packet. Further, a two layer table structure is designed to offer identification of mobile bots and their locations at local WAP networks, and also the global attack information in a central database. Moreover, we noticed that the Equipment Identity (IMEI in short) number possesses a unique feature in cellular communication. Similar to a MAC address of each network adapter, a IMEI number is unique and hard coded to each mobile phone. It nearly impossible to change a IMEI number without somehow damaging the device. Moreover, it is extremely hard to spoof the IMEI numbers because of the cellular communication mechanism [9]. As a result, the IMEI number (rather than the traditional IP addresses in the existing traceback schemes) is used as unique strings for the traceback process. Our analysis and simulation demonstrate that the proposal traceback method is feasible and effective. We believe this work sheds a light on the almost uncharted field of traceback in mobile Internet.

The rest of this letter is organized as follows. We list some preliminaries of the work in Section II, and present the traceback framework in Section III. The feasibility analysis is conducted in Section IV, followed by further discussion in Section V. Finally, we summarize and letter and present the future work in Section VI.

II. PRELIMINARIES

A. IMEI Number

The IMEI number is a 15 decimal digit number, always unique to identify 3 GPP (e.g., GSM, UMTS, and LTE) mobile phones as
well as some satellite phones. Some literatures also describe that this IMEI is 15 or 17 digit number (the additional 2 bits simply indicates software version in device). This unique number to every mobile device (irrespective to smart phone, featured phone etc.) is used in GSM network to identify the location of devices. The term valid IMEI or sometimes genuine IMEI relates to device authentication. It states that the device is legally launched by manufacturer into telecom market. This IMEI number is also hard coded in hardware and making it nearly impossible to change them without somehow damaging the device.

B. Mobile Web Accessing

The framework of WAP (Wireless Application Protocol) application is shown in Fig. 1. In general, a WAP accessing involves two networks, the cellular network and the Internet. Unfortunately, the two networks use different protocols for communication. As a result, a gateway is necessary to bridge the two networks. The gateway is usually called WAP gateway, which typically sits on a telecom carrier network, and perform the specific function of translating plain WML (Wireless Markup Language) in binary WML and binary HTTP into plain text HTTP request. GGSN (Gateway GPRS Support Node) and PDSN (Packet Data Serving Node) are examples of WAP gateway.

It is well known that it is very easy to spoof source IP addresses of IP packets. However, in a cellular network, it is hard to spoof the IMEI number. When a user switched on a mobile device, the SIM card number and the IMEI number respectively along with 128 bit encrypted key (this 128 bit number is stored in SIM card) is transferred to the cellular network to lock a channel for instant and all future communication [10]. When we initiate calls and web browses from mobiles, the cellular network database then checks or authenticates the device based on the transferred IMEI number and also compute the SIM authentication to allow information exchange. Thereafter authentication approval, the cellular network allows the mobile data transfers to cellular and ISP network from mobile through the WAP gateway. In this case it has assumed that the mobile device has a valid IMEI number. It is possible that the IMEI number may be invalid, but in that case the communication will not take place and the user (may be attacker) will be put in gray and black list.

C. Marking On Demand Strategy

Compared with PPM and entropy based traceback methods, DPM is a simple and effective mechanism if the number of attack sources is not that large. However, the number of attack sources to be traced is usually much larger than the capability of the current DPM schemes. In other words, scalability is a critical barrier for DPM based traceback schemes. Yu et al. [5] found a fact that we need to traceback to active bots who are committing attacks, rather than all the bots of a botnet, in practice. As a result, they proposed the marking on demand (MOD) strategy to identify active bots in IP networks. The MOD strategy breaks the scalability constraint of the DPM based traceback schemes, and makes traceback feasible in the Internet. However, the IP address based MOD scheme cannot be hired to traceback mobile bots who are hosted in cellular networks.

III. THE TRACEBACK SCHEME

We confine our traceback under the following conditions in order to make our presentation simple and easy to be understood.

- The attack is detectable at least at the victim side.
- The WAP gateway is not compromised, and the source IP address of the WAP gateway is not spoofed in the IP networks.
- The involved WAP gateways are collaborative in the traceback processing.

As shown in Fig. 2, we set up a global traceback server, which will work as a data center for the traceback purpose. The WAP gateways, the victims and the traceback server will collaborate in the traceback process.

Scalability has been always a problem of DPM based traceback methods. In the proposed framework, we design a two layer table structure to break the scalability constraint. We set up a table at WAP gateways and the global traceback server, respectively. The table at a WAP gateway, gateway table, is defined as follows.

\[
\text{gateway table} = \langle \text{MarkID}, \text{Timestamp}, \text{IMEINumber}, \text{Status} \rangle.
\]

In this table, the MarkID and Timestamp are used to match the feedback from the traceback server. The IMEINumber is used to identify the compromised mobiles in the WAP network by the local network administrators. The status column is used to indicate the state of a suspicious request, which is initially marked as ‘suspicious’, and will be updated if the traceback server confirms the attack to the WAP gateway.

At the traceback server, we have the following table, traceback table, to record global attack information.

\[
\text{traceback table} = \langle \text{WAPIP}, \text{MarkID}, \text{Timestamp}, \text{AttckType} \rangle.
\]
The content of this table populated by the information from the Internet based victims, who detect various attacks, and submit the information to the global traceback server, which will in turn delivers the attack information to related WAP gateways, respectively. This table also offers researchers a global view of various attacks from mobile Internet users.

We note that the Timestamps in the two tables are used to handle the time difference of using the same MarkID as the MOD strategy uses the markable space in a round robin style.

It usually requests sufficient data to successful detection of an attack. For example for DDoS attacks, a WAP gateway may realize a surge of flows which may cause by either a DDoS attack or a genuine flash crowd. We treat this as suspicious flows, and then initiate the marking process.

Combining with Fig. 2, we describe the traceback procedure as follows.

1) When compromised mobiles send attack packets to a WAP gateway, and the gateway noticed that the packets are suspicious, then the gateway will extract the IMEI number from the requests, and map the IMEI numbers with unique MarkIDs, respectively. The MarkIDs are then padded into the marking space of the outgoing IP packets. At the same time, the information of the IMEI numbers, the unique MarkIDs, and the timestamps are deposited to the local gateway table.

2) Once a victim confirms the attack, it extracts the source IP address of the WAP gateway and the unique MarkIDs, and submits the information to the traceback server.

3) The traceback server deposits the information from the victim with the current timestamp, and further notices the related WAP gateway with the unique MarkIDs.

4) The WAP gateways can identify the IMEI numbers against the MarkIDs and the timestamps from the traceback server to identify the compromised mobiles. As a result, the requests from the same mobile to the same Internet server will be blocked. The owners of the mobiles may be noticed via SMS or other means to take actions to remove the bots from the related mobiles.

IV. FEASIBILITY ANALYSIS

In this section, we examine the feasibility of the proposed traceback scheme.

Theoretically, there is no scalability problem at the traceback server end as the table space of the traceback table could be unlimited. As a result, the feasibility depends on individual WAP networks.

First of all, we investigate the sufficiency of marking space of a single IP packet within a WAP network. As we know the available marking space in an IPv4 packet is 16, 17, or 24 bits [3], and we represent the marking space as a random variable $L$. Rajab et al. [11] found that the active bots of a botnet is at the hundreds or a few thousands level. We use random variable $N$ to denote the size of total active mobile bots of a botnet. Here, we consider the worst case for our traceback scheme: all the active bots of an attacking botnet is within one WAP network.

We list the unique MarkID (marking space) that we can have from one IP packet in Table I.

Given that $N$ is around a few thousands level, we can definitely meet the marking space requirement for a single attack in the worst case. Moreover, we can also handle this extreme situation even the size of active bots double or triple in the near future.

Secondly, it is obvious that the bots of an attacking botnet are not behind one WAP network in practice. Moreover, it is possible that multiple botnets hosted by the same WAP networks may carrying out attacks at the same time. We are interested to know whether the proposed traceback method feasible in the practical situations.

Research has demonstrated that malware distribution basically follows power law in terms of networks [12]. Therefore, we employ power law to represent the distribution of active mobile bots in terms of WAP networks in the mobile Internet environment. The power law distribution could be represented by the Zipf distribution, which is defined as follows.

$$p[x = i] = \frac{C}{i^\alpha}, \quad (1)$$

where $\alpha$ is a positive parameter, $p[x = i]$ represents the probability of the $i^{th}$ ($i = 1, 2, \ldots$) largest botnet in terms of size, and $\sum_i p[x = i] = 1$.

Based on equation (1), we obtain

$$C = \frac{1}{\sum_{i=1}^{n} \frac{1}{i^\alpha}}. \quad (2)$$

If we suppose all attacking mobile bots of a botnet come from $n$ WAP networks, and the smallest involved WAP network has just one mobile bot, then the following holds.

$$\frac{C}{n^\alpha} \cdot N = 1. \quad (3)$$

Our study on the large scale data sets of malware distribution offered from reference [13] demonstrates that $\alpha$ is usually around 1.1 for botnets. We therefore use similar values of $\alpha$ in the rest of this analysis.

Based on equations (2) and (3), we can estimate the relationship between the size of an attack botnet and the number of WAP networks involved, and the result is shown in Fig. 3.
Fig. 4. The average number of active bots in a WAP network given the size of the botnet.

Further, we expect to know in average how many mobile bots are hosted by a WAP network. Based on equation (1), we have the mean as follows.

\[ \bar{N} = \sum_{i=1}^{n} P[x = i] \cdot N_i = N \sum_{i=1}^{n} \left( \frac{C_i}{i^\alpha} \right)^2, \tag{4} \]

where \( N_i \) is the size of the bots in the \( i^{th} \) largest WAP network in terms of the number of active mobile bots. Based on the parameters of the previous analysis, we can calculate the mean and the result is shown in Fig. 4.

If we take the average when \( \alpha = 1.1 \), then we have \( \bar{N} = 152 \). We use this as the estimation of the average size of active mobile bots in an attacking botnet, then the concurrent attacks we can handle for the traceback purpose at one WAP network is as follows.

The result in Table II indicates that we can perform the traceback process effectively when there are up to minimum 431 different botnets are carrying out attacks from one WAP network at the same time, which is sufficient to serve the current needs. Therefore, we conclude that the proposed traceback scheme is effective in practice under the current mobile botnet size.

Table II

<table>
<thead>
<tr>
<th>Concurrent Length</th>
<th>Concurrent Attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>L=16</td>
<td>431</td>
</tr>
<tr>
<td>L=17</td>
<td>862</td>
</tr>
<tr>
<td>L=24</td>
<td>110,376</td>
</tr>
</tbody>
</table>

V. FURTHER DISCUSSION

In this letter, we focus on a novel and feasible traceback scheme in mobile Internet environment. The proposed method can be extended in several directions to improve the system in practice. We list a few important ones as follows.

1) The traceback architecture. In this proposal, the traceback architecture is centralized, which inherits many disadvantages of the centralized nature, such as performance bottleneck, reliability, and so on. In practice, we can extend it to a distributed system, such as multiple traceback servers, to address the problem.

2) Scalability. In this letter, we confined our marking space within one IP packet, which is sufficient to deal the current level of cyber attacks. However, we can use multiple IP packets for marking, then the marking space will be dramatically extended for possible usage in the future.

3) Detection at WAP gateways. In the proposed model, it is critical for the WAP gateways to identify suspicious packets. Otherwise, the malicious packets will not be marked, and we cannot traceback to the mobiles although the attack is detected by the victim.

VI. SUMMARY AND FUTURE WORK

In this letter, we proposed a practical traceback method for attack source traceback in mobile Internet environment. The proposed method take the advantage of IMEI number of mobile phones, and employed the marking on demand philosophy. A two layer information storage strategy is designed to complete the traceback purpose. The feasibility analysis indicates that the proposed method is feasible, and can server the current needs of cyber security with a single IP packet.

In regards of future work, many further work can be done in various perspectives, such as make the system distributed, scalability extension of multiple packet marking, performance aware system design, accuracy analysis, and so on.

REFERENCES