Analysis of DNS Security Threats on campus network

Feng Sha*, Ying Wei^, Wenzhi Wu*,
*Network Information Center, Xiamen University of Technology, Xiamen, China
^Department of Computer and Science Xiamen University Tan Kah Kee College, Xiamen, China
Email: shafeng@xmut.edu.cn, sailwy07@xujc.com.cn, wuwenzhi@xmut.edu.cn

Abstract—DNS, as the entrance to Internet application access, is the only way for all business access. Normal services must access services through DNS, and malicious programs also need to be connected through DNS. In this paper, we focus on the analysis of campus network egress data using traffic analysis technology. And we pay attention to DNS-related data, hoping to find useful information about network attack. Since we collected data for about a month (mainly data about DNS data), we found many attacks and a large number of computer viruses. They are hidden in the normal data stream, but can be recognized by DNS data.

1. INTRODUCTION
In the network, The most basic, important and oldest service is DNS[1] (domain name system) . The domain name system is an effective system on the Internet that solves the naming of online machines. Just like visiting friends, you need to know how others go first. When a host wants to access another host on the Internet, you must first know its address. The IP address in TCP/IP is composed of four numbers separated by "." The composition is always not as convenient as the name, so the domain name system is used to manage the correspondence between the name and the IP [2]. Analysis of DNS data can reveal potential cyber-attack behavior in the Internet. The analysis of historical and real-time DNS data will reveal potential security risks in our school, and can explore the development trend of some kind of network virus which cannot be found by traditional antivirus software, and realize the blocking function.

DNS use udp port 53 for commination. So it is easily to be attacks. Nowadays although DNS over TLS, DNS over https came out, traditional udp is still mainstream. Attackers use intelligence domain names for outreach. A large number of network attacks such as worms, mining, botnets, Trojans, backdoors, and remote control are using DNS requests to communicate with the remote control server. So there are many forms of attacks against dns vulnerabilities on the Internet. They are DNS malformed packet attack,

DNS malformed packet attack[3], DNS malformed packet attack is a kind of network DDOS attack. The attacker sends a defective request packet to the target DNS server, causing DNS to crash or crash while processing such data packets.

Reflection amplification attack[4], the attacker uses the feature that the DNS reply packet is larger than the request packet to amplify the traffic. The source IP of the forged request packet is the victim IP, and the traffic of the reply packet is introduced into the victim server, which degrades the server performance.
Wildcard domain name attack, the attacker initiates a large number of unresponsive domain names to the recursive server, causing the recursive server to continuously consume system resources, thereby affecting normal domain name resolution.

Source address DDoS attack, this type of attack means that a specific infected terminal initiates a large number of illegal requests and continuously consumes recursive server resources, affecting normal resolution.

Target domain DDoS attack\(^5\), the attacker uses the infected terminal or zombie program to initiate a large number of random subdomain requests to the victim's domain to implement a DDoS attack on the authoritative server and related recursive servers in the target domain.

Illegal domain name hijacking attack\(^6\), Domain name hijacking is a method of Internet attack. By attacking the domain name resolution server (DNS) or forging the domain name resolution server (DNS), the target website domain name is resolved to the wrong IP address to achieve the purpose that the user cannot access the target website or The purpose of deliberately or maliciously requiring users to access the specified IP address (website).

DNS cache poisoning\(^7\), DNS cache poisoning refers to that attackers use counterfeit data to pollute the DNS cache, enticing users to access illegal addresses in the DNS cache, to achieve data theft and other attacks.

DNS tunneling attack\(^8\), Attackers use DNS as a covert communication channel to transmit sensitive data outside the network, or remotely control the trapped terminal through a DNS tunnel to carry out other attacks. It can create a C&C tunnel through the DNS protocol, making the attacker more covert. You can access any data and upload and download files, and get a shell. For example, when you enter a network that requires authentication, you find that all IPs and ports are inaccessible, but DNS port 53 can be used, which provides a convenient condition for DNS tunnel attacks.

2. RESEARCH METHOD
In order to deal with the increasingly severe and diversified DNS attack methods, we must maintain the leading technology. And the first step is to correctly identify and discover these attacks. There are two ways in data recognition, payload analysis and traffic analysis.

Analysis of attack principle based on application layer domain name. DNS data implementation attack detection method based on application layer domain name This project mainly adopts the following two types: First, based on the detection method of DNS access traffic change, the change of access traffic of the campus network, the detection error caused by the fluctuation of the DNS access traffic can be effectively reduced. By monitoring the core network nodes of the campus network, the DNS data is mainly monitored in this project. Extract the attacked address and the attack source address by using the access traffic data and Netflow information. Second, based on the DNS attack detection method of the number and distribution of the domain name data request source, in order to hide the attack, the attacker on the internet can reduce the attack rate and make the attack traffic rate close to the normal access rate, thereby increasing the detection difficulty, but DDOS or computer virus attack, the large increase in the number of access to a domain name is a distinct feature of the attack and this feature cannot be hidden. Based on this feature, the analysis is performed using open source lightweight big data analysis technology.

3. TECHNICAL ROADMAP
We plan the following according to the needs of campus network security construction and the amount of log data analyzed. And formulate a reasonable experiment plan for the campus environment of our school.

We first establish an experimental environment.[9,10,11] The experimental environment of this project is established in a pure campus network environment. We have deployed 4 ZDNS devices in our campus network. The dns58 and dns210 is the authoritative DNS, and the ipv6dns is use for ipv6 purpose, the node master is used to control the data synchronization of all nodes. So the ipv4 client can
share the load through dns58 and dns210, ipv6 client can query DNS authoritative data through ipv6dns. The figure1 below is the system diagram.

![Figure1. DNS System in XMUT](image1.png)

We deployed a DNS traffic analysis device on the campus network in bypass mode. For functional testing, we use the command tools nslookup and dig. And for performance testing, we use tool software dnsperf and queryperf. The deployment is as follows figure2:

![Figure2. System architecture diagram](image2.png)

DNS data analysis: The Figure3 and Table1 shows the data analysis of the DNS system currently running by Xiamen University of Technology. We took one mouth DNS data from 10th May to 9th June for our project. There are total 432,458,258 records in the data sheet. From the analysis type, it can be found that 64.05% of the current data query volume is concentrated in the A record, that is, access to the IPv4 network application. The AAAA record, that is, the query volume of the IPv6 network is only 33.91%. Although the proportion of other analysis types is not high, there are still nearly 1 million records such as PTR, SRV, MX, RRSIG, NS, TXT, SOA, CNAME etc.

![Figure3. DNS resolution type statistics in XMUT](image3.png)
DNS uses the A record to record the domain name corresponding to the host name. The AAAA record is used to resolve the domain name to the IPv6 address. The CNAME record is usually an alias resolution. The NS record indicates the domain name server address, and the MX record is usually used for the email system. The TXT record represents a description of a certain host name or domain name, and the PTR record is a reverse DNS resolution. The following table is the classification ranking of our school's request for parsing data within one month. From the table, we can see that the most used requests in the campus network are A records and AAAA records, accounting for about 98% of the total requests.

| Type   | Number   | %    |
|--------|----------|------|
| A      | 276990218| 64.05|
| AAAA   | 146654672| 33.91|
| PTR    | 7599450  | 1.76 |
| SRV    | 567034   | 0.13 |
| MX     | 205665   | 0.05 |
| RRSIG  | 173767   | 0.04 |
| NS     | 91024    | 0.02 |
| TXT    | 63182    | 0.01 |
| SOA    | 57163    | 0.01 |
| CNAME  | 37797    | 0.01 |

Data preprocessing, we transfer the mirrored traffic to the log server and give it to the data analysis software for processing. The DNS data log is huge. Take our campus as an example. The amount of log analysis data for one day is near to 26G now. The processed DNS log data is subsequently processed using the open source software Elasticsearch software\(^{[10]}\). The data extracted from DNS data, and the original modeling tool data needs to be modified. The data is statistically analyzed and analyzed, and the data access in the time period is counted, and the data in the time period of the source IP access is matched according to the access source IP obtained from the information center. Draw a trend chart or curve to explore the relationship\(^{[12]}\). The following figure 4 shows DNS resolution data, taken from the Campus DNS server.

![Figure 4. DNS resolution log format in XMUT](image)

Big data processing, the analysis of DNS query logs using open source software Elasticsearch, Elasticsearch is a distributed search and analysis engine based on the open source Lucene lightweight full-text indexing engine tool. When log server collects the requested domain name, it sends to ES software. And system will count each letter and calculate the probability of each letter appearing, and then combine the Shannon entropy formula to calculate the information entropy of the domain name. The average entropy value is calculated based on the TOP normal domain names commonly used on the Internet to determine abnormal values.

We also need to consider other parameters, such as request time, source IP, destination IP, access type, and access domain name. Reasonably set the weight of each parameter to do in-depth data mining. For example, if a request mostly occurs in the evening after-hours, it is an abnormal access, or a request frequently accesses a common URL, these should be paid attention to.
4. RESULTS AND DISCUSSION

4.1. DNS data security threat analysis

There are many kinds of data related to our DNS. Such as: security data, system data, network data, end user data, and so on. Various data information requires integrated analysis, data behavior analysis, and so on. In the past month as of June 10, the real statistics showed the main threat letters in the campus Network are the following:

Among them, there are 33 alarm threat hosts in the network, 15 alarm events, and the number of threat access initiation is 89707. The main types of threats in the network are: Imwee download Trojan (32.34%), CoinMiner mining Trojan (31.58% of the total network), Ramnit worm (17.06%), and other types of threats (19.02%). 9 high-threat events are: CoinMiner mining trojan, Ramnit worm, NrsMiner mining botnet, Nitol botnet, Palevo botnet, Zegost remote control, Sality infection virus, Dofloit Trojan. 11 middle-level threat incidents were: Imwee downloader Trojan, Ramnit worm, Sality infection virus, Nitol botnet, Incident_1710_Jsctrl, Trojan horse, Diplugem Trojan, XMRCoinMiner mining Trojan, CoinMiner mining Trojan, AutoRun worm.

Types of threats rankings, we mainly listed the proportion of several security threats in the overall threat, as showed in the following table2:

| Rank | Threat                             | %     |
|------|------------------------------------|-------|
| 1    | Imwee download Trojan              | 32.34 |
| 2    | CoinMiner mining Trojan            | 31.58 |
| 3    | others                             | 19.02 |

Threat host addresses are: 58.*.*.222 (number of detections: 32233), 172.*.170.56 (number of detections: 17465), 172.*.48.60 (number of detections: 16959), 210.*.213.3 (Number of detections: 16385), 10.*.8.220 (Number of detections: 3580), 172.*.240.104 (Number of detections: 794), 172.*.232.40 (Number of detections: 784), 172.*.51.180 (Number of detections : 672), 172.*.240.189 (number of detections: 161), 172.*.227.122 (number of detections: 147), 172.*.249.19 (number of detections: 125), 10.*.94.190 (number of detections: 113), 10.*.66.194 (Number of detections: 105), 10.*.7.39 (Number of detections: 91), 120.*.98.236 (Number of detections: 90), 120.*.99.234 (Number of detections: 30).

IP ranking, we list the host IP that have threatened access in the campus network, as showed in the following table2:

| Rank | IP       | Quantity |
|------|----------|----------|
| 1    | 58.*.*.222 | 32233    |
| 2    | 172.*.170.56 | 17465    |
| 3    | 172.*.48.60 | 16959    |
| 4    | 210.*.213.3 | 16385    |
| 5    | 10.*.8.220  | 3580     |
| 6    | 172.*.240.104 | 794      |
| ......| ......     | ......    |
| 16   | 120.*.99.234 | 95       |

From the above host. Statistics of domain names IP initiates the analysis of the overall ranking of security threats; the order of IP requests is arranged in order, which can help the relevant O&M personnel to clearly understand the IP Internet access trajectory of each end user and whether there are attacks. Therefore, we should check whether the IP "172.*.170.56" has an offensive behavior.
4.2. Future discussion

There are two suggested problems. A key issue is the threat of accurate source tracing of the host. This project is to directly monitor the traffic bypass through the network exit mirror. In our analysis, we found that the authoritative DNS address of our school would be mistaken as a threat IP. The reason is that the client computer initiates access through the authoritative DNS of our school, and the authoritative DNS forwards the request. The system mistakenly thought it was an authoritative DNS attack threat. Therefore, we must combine the logs of authoritative DNS for matching analysis to accurately trace the source of the IP and find the real attacker.

The other one is that security threats cannot be blocked in time. Our analysis of security threat data is an after-the-fact audit. When we find a threat, a security threat has occurred or has existed for a while. The relevant system may have been attacked. The most effective means is to stop the attack while discovering it. This requires linkage with devices that have blocking capabilities such as firewalls, IPS.

5. CONCLUSION

It is clear that the analysis of the campus DNS data has a very large analytical value. We can find out the all network running status and discover network attack behaviors immediately such as DDOS, computer virus or other threat. We analyzed a month of DNS data on the campus network and found a certain number of security threats and attacks against DNS. However, the current attack methods against DNS are becoming more and more covert, and it is necessary to study a deeper level of detection methods to cope with the increasingly serious DNS security threats.

ACKNOWLEDGMENT

Fund Project: This research was support by CERNET Innovation Project (NGII20180510). And also support by ZDNS.

REFERENCES

[1] Albitz, DNS and BIND[M]. 2001.
[2] Almond C. DNSSEC validation and BIND9 cache[J]. 2014.
[3] Magdych J S , Rahmanovic T , Mcdonald J R , et al. Secure gateway with firewall and intrusion detection capabilities[J]. 2012.
[4] T Liu C, Xiong G, Liu J, et al. Detect the reflection amplification attack based on UDP protocol[C] // 2015 10th International Conference on Communications and Networking in China (ChinaCom). IEEE, 2016.
[5] Jie Z. DDOS attack source tracing algorithm based on the IP address re-stitching[J]. computer era, 2012.
[6] Borgwart A, Boukoros S, Shulman H, et al. Detection and Forensics of Domains Hijacking[C] // IEEE Global Communications. IEEE, 2015.
[7] Sooel Son, Vitaly Shmatikov. The Hitchhiker's Guide to DNS Cache Poisoning[C] // Security & Privacy in Communication Networks-ternational Icst Conference. DBLP, 2010.
[8] Franz F. Esberger. DNS Tunneling[J]. 2010.
[9] Deleuze G. Kafka: Toward a minor literature[M]. U of Minnesota Press, 1986.
[10] Gormley C, Tong Z. Elasticsearch: The definitive guide: A distributed real-time search and analytics engine[M]. " O'Reilly Media, Inc.", 2015.
[11] Bagnulo M, Matthews P, van Beijnum I. Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers[R]. 2011.
[12] Ruizhe Ye, Ying Wei, Wenzhi Wu, FengSha, DNS network security analysis based on Campus Environment[J]. ICIS2019.