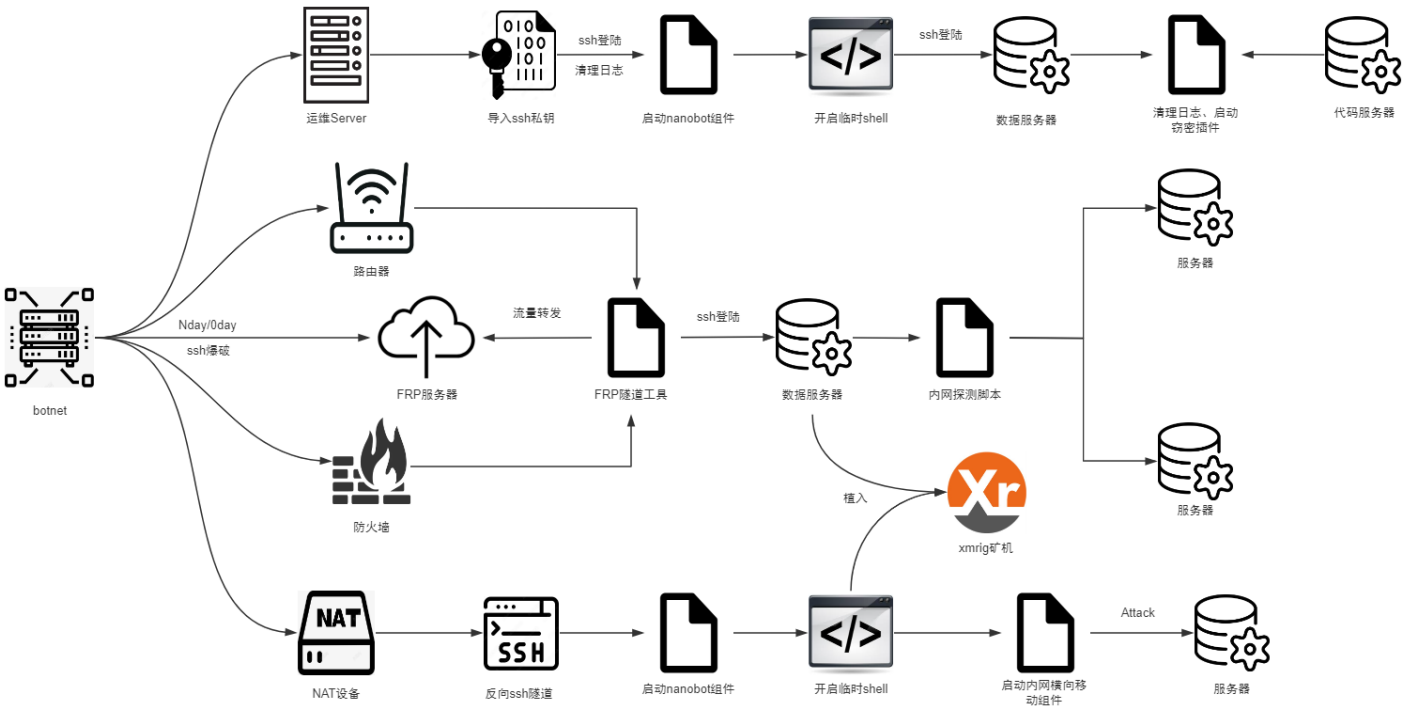
Operation Veles: Decade-Long Espionage Targeting the Global Research and Education Sector

Overview

For a long time, security vendors have had limited coverage in researching espionage incidents related to Linux systems. Many disclosed APT attacks focused on office machines (i.e. Windows platforms), result in data theft primarily involving non-sensitive internal documents. We believe that this type of espionage attack garners more attention than the actual harm it poses. However, in the field of scientific research, Linux servers often host critical data, making their security of utmost importance. Therefore, strengthening security research and event monitoring for Linux systems is a crucial task in safeguarding the high-quality development of national science and technology.

Since the initiation of “UTG” (Unknown Threat Group) numbering, QiAnXin Threat Intelligence Center has closely monitored attacks targeting server environments within government and enterprise sectors, and have discovered several threat actors such as UTG-Q-008 and UTG-Q-009, which have caused significant harm to government and enterprise entities. Among them, UTG-Q-008 is the only threat group exclusively targeting Linux platforms for its malicious activities. After a year-long intensive tracking effort, we have finally confirmed evidence of UTG-Q-008 utilizing the resources of a massive botnet network for espionage activities against the domestic research and education sector. Up to 70% of the infrastructure are springboard servers, with a different batch of springboard servers being used for each new activity. The domain names controlled by the attackers have been active for at least a decade, displaying an adversarial strength far surpassing mainstream APT groups, which has deeply impressed upon us the notion that huge network resources are the best weapons. The attack flowchart is shown below:



Target

UTG-Q-008 has multiple attack lists for its domestic activities, and we have obtained one of them, which includes over five thousand domestic network segments:

```
Prefix /12: 6 networks, 6291456 total IPs
Prefix /13: 12 networks, 6291456 total IPs
Prefix /14: 15 networks, 3932160 total IPs
Prefix /15: 45 networks, 5898240 total IPs
Prefix /16: 86 networks, 5636096 total IPs
Prefix /17: 72 networks, 2359296 total IPs
Prefix /18: 107 networks, 1753088 total IPs
Prefix /19: 303 networks, 2482176 total IPs
Prefix /20: 1252 networks, 5128192 total IPs
Prefix /21: 671 networks, 1374208 total IPs
Prefix /22: 493 networks, 504832 total IPs
Prefix /23: 594 networks, 304128 total IPs
Prefix /24: 1553 networks, 397568 total IPs
Prefix /25: 2 networks, 256 total IPs
去重后的总IP数量: 17172070
```

After deduplication, it contains over 17 million target IP addresses within China. Detailed comparisons have confirmed that the majority of these target IPs belong to the CN CER (China Education and Research) network assets, displaying a high level of specificity. In addition to this, UTG-Q-008 has shown a strong interest in top-tier biological genetics and RNA immunotherapy research projects in both China and the United States. Threat groups with such interests in this field are rare, and similar attack activities are found in Operation HideBear [1].

Botnet

From the defender's perspective, attackers have virtually unlimited network resources. Each time a large-scale operation is launched, the domain names for payload requests and the IP addresses for shell bouncing on victim Linux servers are brand-new springboard servers. The attack activities typically occur between 0-4 a.m., and the duration of the shell is only 2-3 minutes. Short duration of shell renders traditional IOC (indicators of compromise) intelligence ineffective in defending against them.

Scanning and Brute-forcing

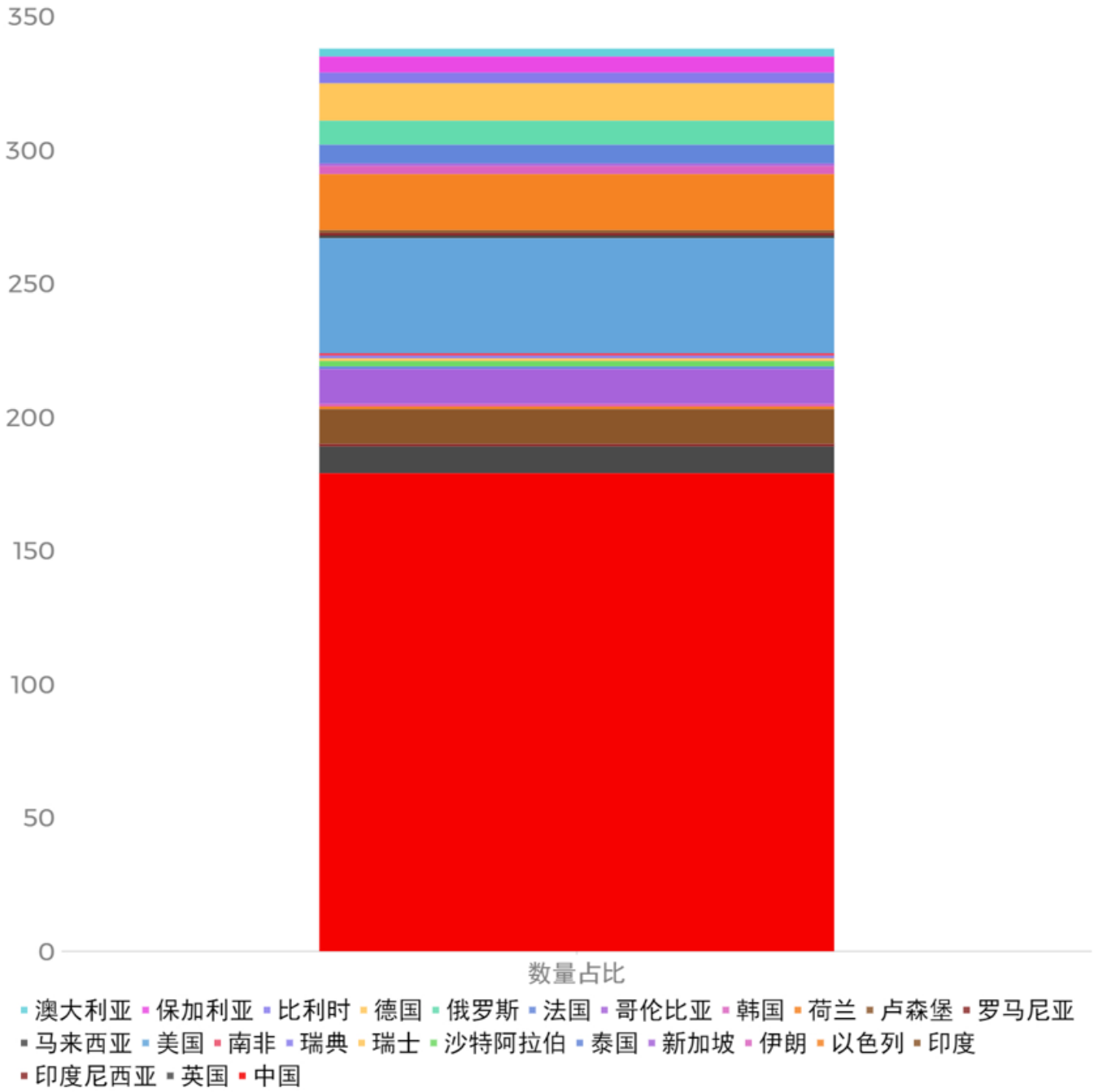
Currently, many organizations don’t use default SSH ports any more on their Linux servers located at the network perimeter. Therefore, UTG-Q-008's first step is to utilize the massive network resources of botnet to perform distributed SYN scans to identify open ports on the target networks. We calculated the SYN scan frequency for individual IP addresses, averaging 25-35 scans per second. Similarly, in subsequent

distributed brute-forcing activities, the number of brute-force attempts per second from a single IP does not exceed ten. Under this adversarial strategy, UTG-Q-008 has managed to create a small brute-forcing footprint. Within a month, they successfully brute-forced the root passwords of nine servers, including six research servers and three perimeter devices, mainly firewalls, routers, and out-of-band management for hosts.

In our long-term engagement in tracking targeted attacks, this is the first time we observed direct involvement of a botnet in espionage. The scale and quality of the affected entities have exceeded our expectations. In previous APT cases, achieving such "impressive results" in the Linux server domain would not be possible without a few 0-days.

Distribution of Botnet Resources

We conducted a simple analysis of the quantity and geographical distribution of the source IP addresses used for SSH logins. There are the highest number of controlled nodes in China, followed by the United States.



There is no significant uniform characteristic among the hundreds of controlled nodes. Only a few dozen nodes host web servers with Zabbix PowerMTA monitoring. During the rollback process in QiAnXin's botnet monitoring system, we discovered three nodes associated with the Perlbot botnet, three nodes associated with Outlaw, and one node was linked to the Mirai botnet. The Nanobot released by the attackers during lateral movement is very similar to Perlbot. Since Perlbot itself is a simple script Trojan that anyone can use, we can only confirm that the botnet network resources targeting domestic entities can be accessed by UTG-Q-008. We cannot attribute the Mirai nodes' cluster to a specific entity, since the occurrence of overlap between nodes from two different botnet is considered normal.

The involvement of botnets in espionage activities is not uncommon. The key lies in the extent of their participation. For example, in 2024, the Moobot botnet provided network proxies to APT28 for spear-phishing email delivery [2]. In 2019, Lazarus utilized the TrickBot botnet to distribute exclusive malware for attack activities [3]. However, based on our a-year-long tracking of UTG-Q-008, we believe that the botnet

behind this threat group is directly involved in espionage activities, because it’s deeply engaged in the aspects from target reconnaissance, brute-forcing, vulnerability exploitation, Trojan components delivery to C&C infrastructure. Now, how does it perform on the Windows platform?

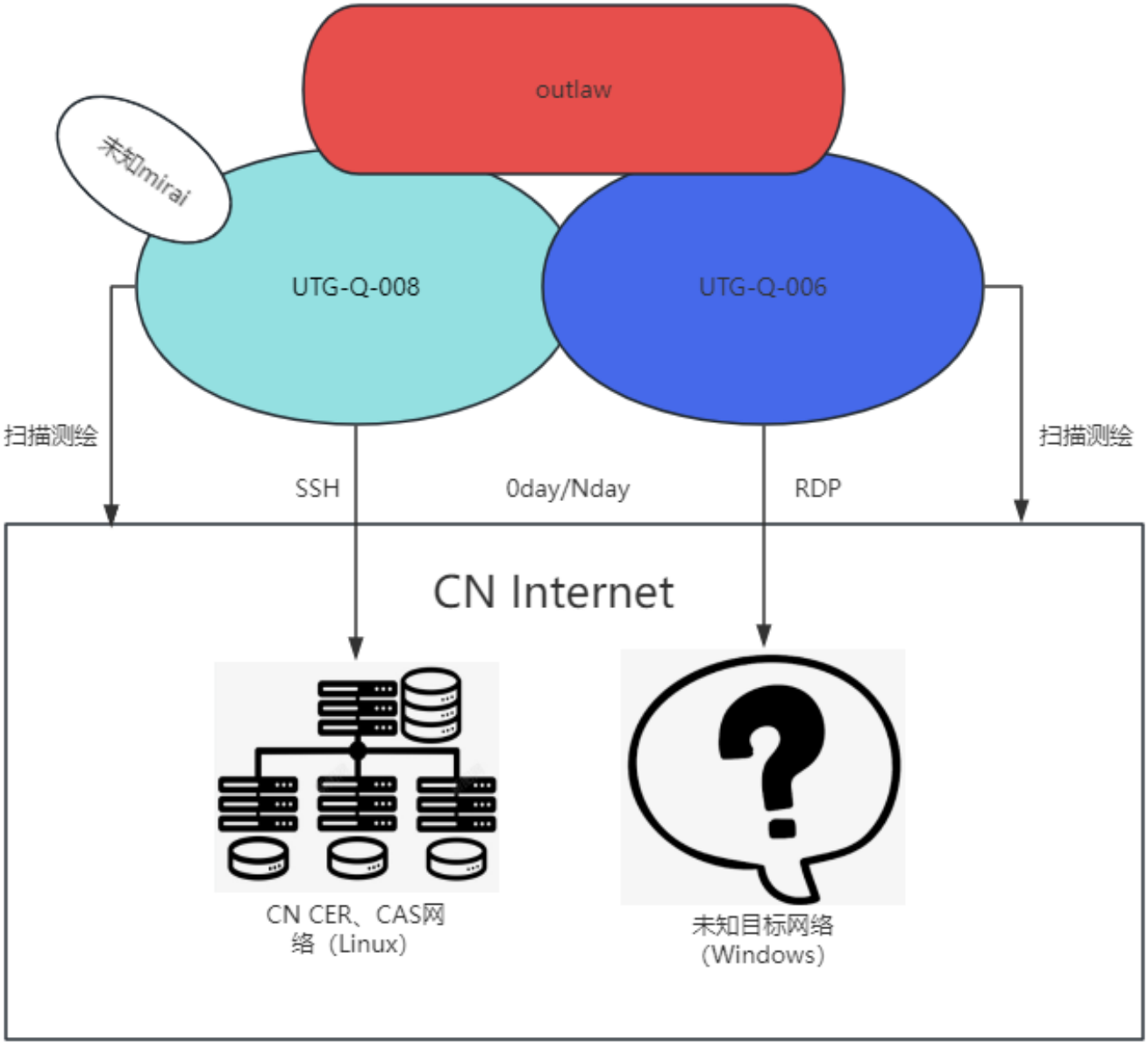
Connection Between UTG-Q-006 and UTG-Q-008

UTG-Q-006 is a threat group that primarily targets Windows devices exposed on the public Internet. It also relies on a large botnet. Although we don't have the target list for UTG-Q-006, attackers managed to successfully brute-force the RDP (Remote Desktop Protocol) ports of critical entities within half a month using eight-character non-weak passwords. During lateral movement, UTG-Q-006 demonstrated sophisticated LOLbins techniques, roaming within the internal network using legitimate tools such as AnyDesk, Chisel, and Advanced Port Scanner, ultimately infiltrating the MES (Manufacturing Execution System) server. This activity poses potential implications for industrial production processes.

We compared the brute-forcing nodes of UTG-Q-006 with those of UTG-Q-008 and discovered several overlapping nodes. Furthermore, it is noteworthy that several hundred brute-forcing nodes controlled by UTG-Q-006 also exhibited activity related to dozens of Outlaw botnet nodes within the botnet monitoring system.

```
cd ~ && rm -rf .ssh && mkdir .ssh && echo "ssh-rsa
AAAAB3NzaC1yc2EAAAABJQAAAQEArDp4cun2lhr4KUhBGE7VvAcwdli2a8dbnrTOrbMz1+5O73fcBOx8NVbUT0bUanUV9tJ2/9p7+vD0Epz3Tz/+0kX34uAx1RV/75GVOmNx+9EuWOnvNoaJe0QXxziIg9e
TDengvS8hX1kNFS4Mjux0hJOK8rvcEmPecjdySYMb66nylAKGwCEE6WEQHmd1mUPgHwGQ0hWCwsQk13yCGPK5w6hYp5zYkFnvlC8hGmd4Ww+u97k6pfTGTUbJk14ujvcD9iUKQTTWYYjIu5PmUux5bsZ0F
Rw== mdrfckr">>.ssh/authorized_keys && chmod -R go= ~/.ssh && cd ~
```

Due to the complexity of botnets, we can only confirm the overlap between UTG-Q-008, UTG-Q-006, and the Outlaw network. We cannot determine whether they have an employer-employee relationship or a hierarchical relationship. The overlap of nodes from different attack groups is as follows:



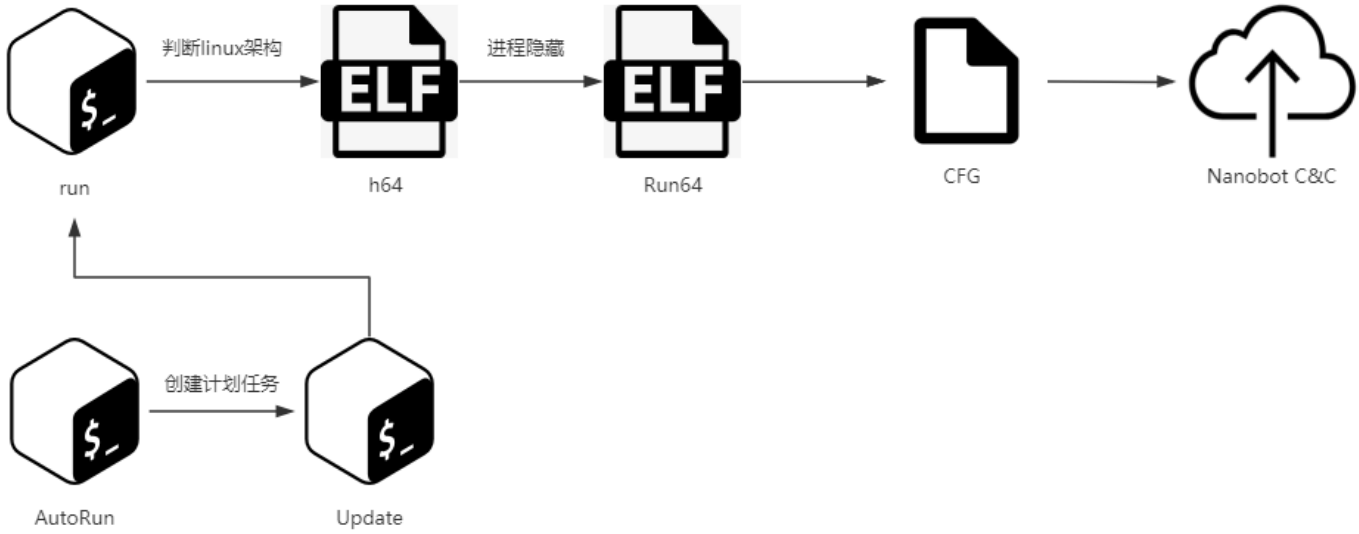
During our monitoring period over a year, these botnet nodes targeting domestic entities have never initiated any DDoS activities, which is highly unusual for traditional botnets.

Weapon Components

UTG-Q-008's weapons are usually packaged in tar format and stored on the springboard servers. The infrastructure does not overlap with the aforementioned botnet nodes. The services running on the compromised servers are mostly disorganized, with only the WordPress framework being identifiable. Additionally, most of the springboard servers have domain names, including a legitimate domain name in China that has been active for 14 years. Therefore, attackers typically operate using domains of springboard servers. The distribution of springboard servers by country is as follows:

Nanobot

Once the attackers gains control of a server, they typically download the Nanobot component from the springboard server using wget or CURL. The startup process is as follows:



The attacker refers to the executed Run64 in the comments as Nanobot. After analyzing the ELF executable, it was determined that it is packaged with Python and has a core logic very similar to the open-source Perlbot.

From the continuous network traffic, we can confirm that once the Nanobot establishes a C2 (Command and Control) connection, the attacker chooses to initiate new reverse shells or SSH reverse tunnels to download subsequent plugins. These temporary shells connecting to the springboard server C2 do not overlap with the aforementioned botnet nodes or the springboard servers storing weapons. Furthermore, the shells only persist for 2-3 minutes, making it difficult to capture and analyze them. The distribution of the springboard server IP types for the reverse shells includes Ubiquiti routers, unknown smart home devices, exchange servers, etc.

Internal Network Detection Component

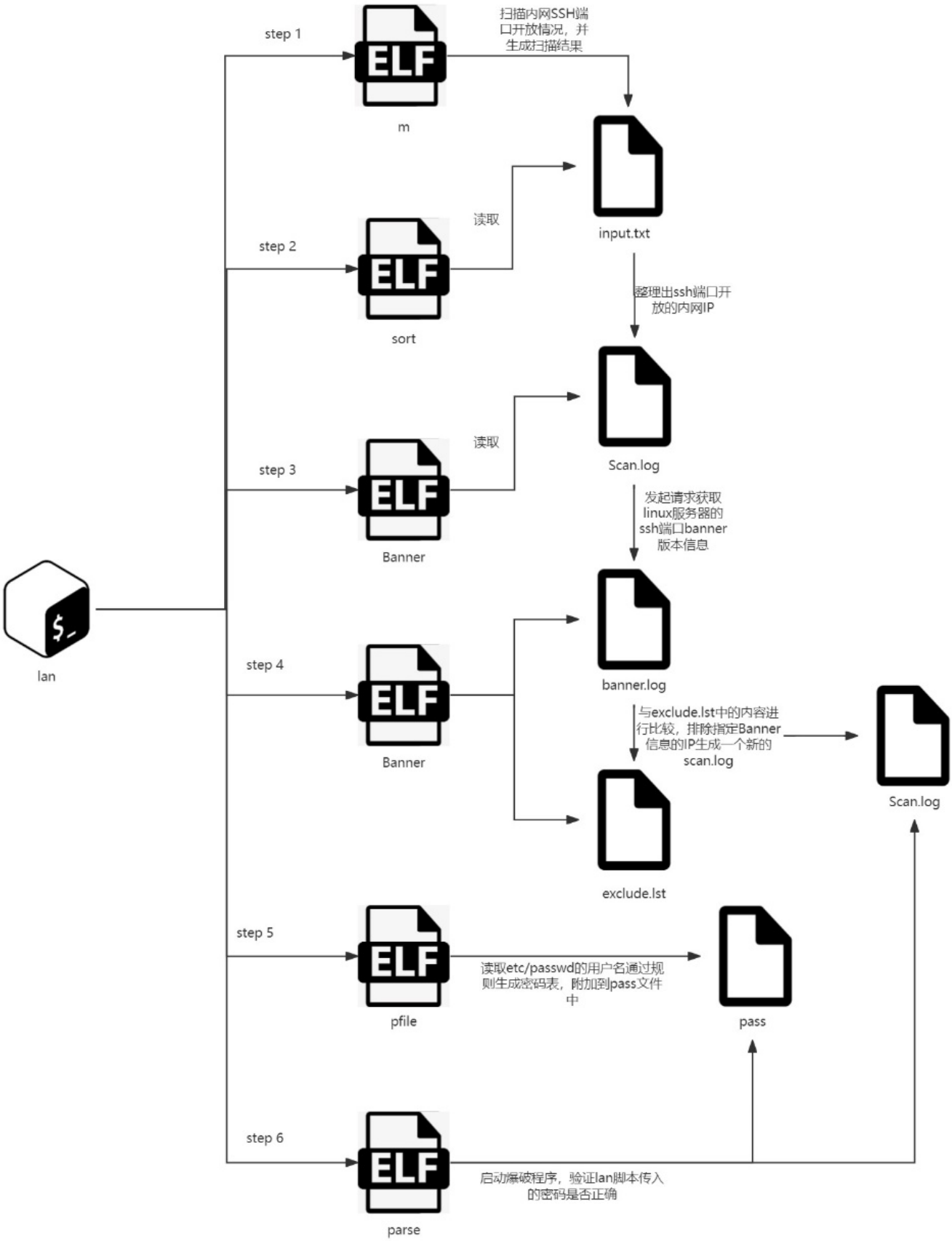
UTG-Q-008 possesses multiple types of internal network scanners, typically used to scan designated ports on machines within the B segment of the internal network.

```
49   printf("scanning: %s", argv[1]);
50   fflush(stdout);
51   memset(v16, 0, 0x100uLL);
52   init_sockets();
53   v13 = time(0LL);
54   while ( 1 )
55   {
56     if ( v6 )
57     {
58       v4 = tot;
59       v5 = time(0LL);
```

Once the attacker gathers the network segments within the internal network, they deploy lateral movement components.

Lateral Movement Component

The process of deploying lateral movement component is as follows:



The overall process consists of two stages. In the first stage, based on the results generated by the B segment scanner, attackers grabs SSH port banner on the target's Linux servers. They compare the results “banner.log” with the built-in “exclude.lst” file in the toolkit to exclude those Linux servers with

specific SSH banners from next steps. The attacker's brute-forcing program may be optimized for certain SSH versions.

```
SSH-2.0-dropbear_2018.76
SSH-2.0-dropbear_2019.78
SSH-2.0-dropbear_2020.80
SSH-2.0-dropbear_2020.81
SSH-2.0-dropbear_2022.82
SSH-2.0-FreSSH.0.8
SSH-2.0-Gitblit_v1.9.3 (SSHD-CORE-1.2.0-NIO2)
SSH-2.0-Go
SSH-2.0-HUAWEI-1.5
SSH-2.0-HUAWEI-UMG8900
SSH-2.0-HUAWEI-VRP-3.10
SSH-2.0-HUAWEI_VRPV8.0
SSH-2.0-IPSSH_5.1.0p1
SSH-2.0-IPSSH-6.9.0
SSH-2.0-lancom
SSH-2.0-libssh-0.6.0
SSH-2.0-libssh_0.8.0
SSH-2.0-Mocana SSH
SSH-2.0-Mocana SSH 5.3.1
SSH-2.0-mpSSH_0.0.1
SSH-2.0-NetScreen
SSH-2.0-paramiko_2.12.0
SSH-2.0-paramiko_2.1.3 501 command not implemented ERROR
```

The main function of the second stage is to use “pfile” program to read the /etc/passwd file of the Linux servers, retrieve usernames, and generate additional password dictionaries by appending weak passwords to the usernames. For example, "root+1234". The newly generated password dictionary is added to the built-in “pass” file in the toolkit.

```
.text:0000000000490400
.text:0000000000490400 lea     r12, [rsp+var_660]
.text:0000000000490408 cmp     r12, [r14+10h]
.text:000000000049040C jbe     loc_491674
.text:000000000049040C
.text:0000000000490412 sub     rsp, 6E0h
.text:0000000000490419 mov     [rsp+6E0h+var_8], rbp
.text:0000000000490421 lea     rbp, [rsp+6E0h+var_8]
.text:0000000000490429 lea     rax, aEtcPasswd                 ; "/etc/passwd"
.text:0000000000490430 mov     ebx, 0Bh
.text:0000000000490435 xor     ecx, ecx
.text:0000000000490437 xor     edi, edi
.text:0000000000490439 call    os_OpenFile
.text:0000000000490439
.text:000000000049043E xchg    ax, ax
.text:0000000000490440 test    rbx, rbx
.text:0000000000490443 jz      short loc_4904C3
.text:0000000000490443
.text:0000000000490445 movups  [rsp+6E0h+var_2C8], xmm15
.text:000000000049044E movups  [rsp+6E0h+var_2B8], xmm15
.text:0000000000490457 lea     rdx, unk_499660
.text:000000000049045E mov     qword ptr [rsp+6E0h+var_2C8], rdx
.text:0000000000490466 lea     rdx, off_4CA188                 ; "Error opening /etc/passwd:"
.text:000000000049046D mov     qword ptr [rsp+6E0h+var_2C8+8], rdx
.text:0000000000490475 jz      short loc_49047B
.text:0000000000490475
.text:0000000000490477 mov     rbx, [rbx+8]
.text:0000000000490477
.text:000000000049047B
.text:000000000049047B loc_49047B:                             ; CODE XREF: main_main+75↑j
.text:000000000049047B mov     qword ptr [rsp+6E0h+var_2B8], rbx
.text:0000000000490483 mov     qword ptr [rsp+6E0h+var_2B8+8], rcx
.text:000000000049048B mov     rbx, cs:os_Stdout
.text:0000000000490492 lea     rax, go_itab__ptr_os_File_comma_io_Writer
```

The original “pass” file is a password dictionary specifically designed by UTG-Q-008 for targets in China, containing over 4,000 account-password combinations, with the majority being based on Chinese Pinyin.

```
1  wangrui wangrui
2  lingfengguo glf123
3  zhangxinkui 666666
4  root honor_wuzhen@Sugon;2022
5  user Huawei12#$
6  rizhiyi rizhiyi&2014
7  root wanzai@2021
8  root www.163.com
9  sugon fdnh3#RF|##@|RDFC
```

After a detailed analysis, we believe these account-password combinations are not randomly generated but accumulated by the attacker over many years of attack activities in China. It is conservatively estimated that thousands of servers in China have been compromised in UTG-Q-008's historical activities. The “parse” program is then initiated to start brute-forcing the internal network servers. First, an HTTP request is made to read data from the built-in springboard server URL to validate the parameters passed in the “lan” script. The brute-forcing process only proceeds after successful network validation.

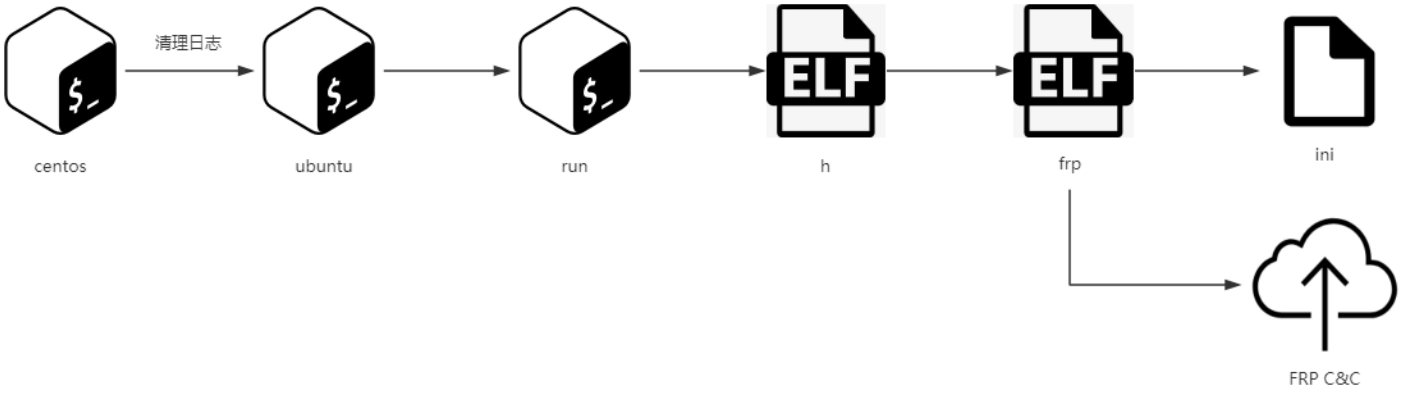
The attacker may find the six ELF executables written in Golang in the toolkit too cumbersome, so they released a lightweight Python script during lateral movement on other machines. This script abandons the SSH banner grabbing process and retains only the logic of the second stage.

```
        username + username + "1234",
        username + username + "12345",
        username + username + "123456",
        username + "1!",
        username + "@00",
        "cvlab" + username + "123!",
        "ustc-" + username + "-123",
        username + "0212" + username,
        "korea2022",
        "korea2023",
        "ubuntu"
    ]

    with open("P", "w") as password_file:
        for password in fixed_passwords:
            password_file.write(f"{username}:{password}\n")

    all_passwords = fixed_passwords

    # Use a pool of worker threads for parallel execution
    with ThreadPoolExecutor() as executor:
        results = list(executor.map(su_try_pwd, [username] * len(all_passwords), all_passwords))

    busted_passwords = set(password for result, password in results if result)
```

Due to the years of effort that UTG-Q-008 put in China, this component designed by them has achieved remarkable results in lateral movement within Linux server networks.

FRP Component

When the attacker wants to access machines within the internal network, they generally start the FRP reverse proxy on the boundary server. The execution process is as follows:



FRP also serves another purpose: when the lateral movement component within the internal network has limited effectiveness, FRP can be used to leverage the computing power of external botnets to brute force critical machines within the internal network.

Espionage Plugin

After infiltrating the internal network to a certain extent, the attacker will choose to install an espionage plugin on important servers. When the ELF is executed, it runs an embedded bash script. The bash file contains numerous regular expressions used to collect sensitive information stored on Linux servers. The functionality of these regular expressions can be divided into several parts, each with around ten matching rules. We will introduce some selected rules.

1. Parsing various historical and hidden files on Linux to identify potential commands containing sensitive information, such as SSH and FTP.
2. Searching for VNC credentials and downloading decryption plugins for decryption.
3. Extracting credentials from sshpass, GitHub, and other types of credential files.
4. Searching for multiple files that meet certain criteria and using corresponding regular expressions to extract sensitive information from the files. For example, searching for .gitconfig files in specific directories to extract email information.
5. Analyzing system logs and code, such as searching for plaintext account-password combinations in files under the /usr/include/ directory.

During the analysis, it was discovered that UTG-Q-008 has implemented independent filtering conditions for files under the "postech" directory with a size of less than 15k. Through further investigation, "postech" seems to refer to a research-oriented university in South Korea. It is unclear why this rule is used among activities targeting China's research system. It could be that the attackers forgot to remove it or they wanted to obtain information on cooperative projects between "postech" and domestic entities.

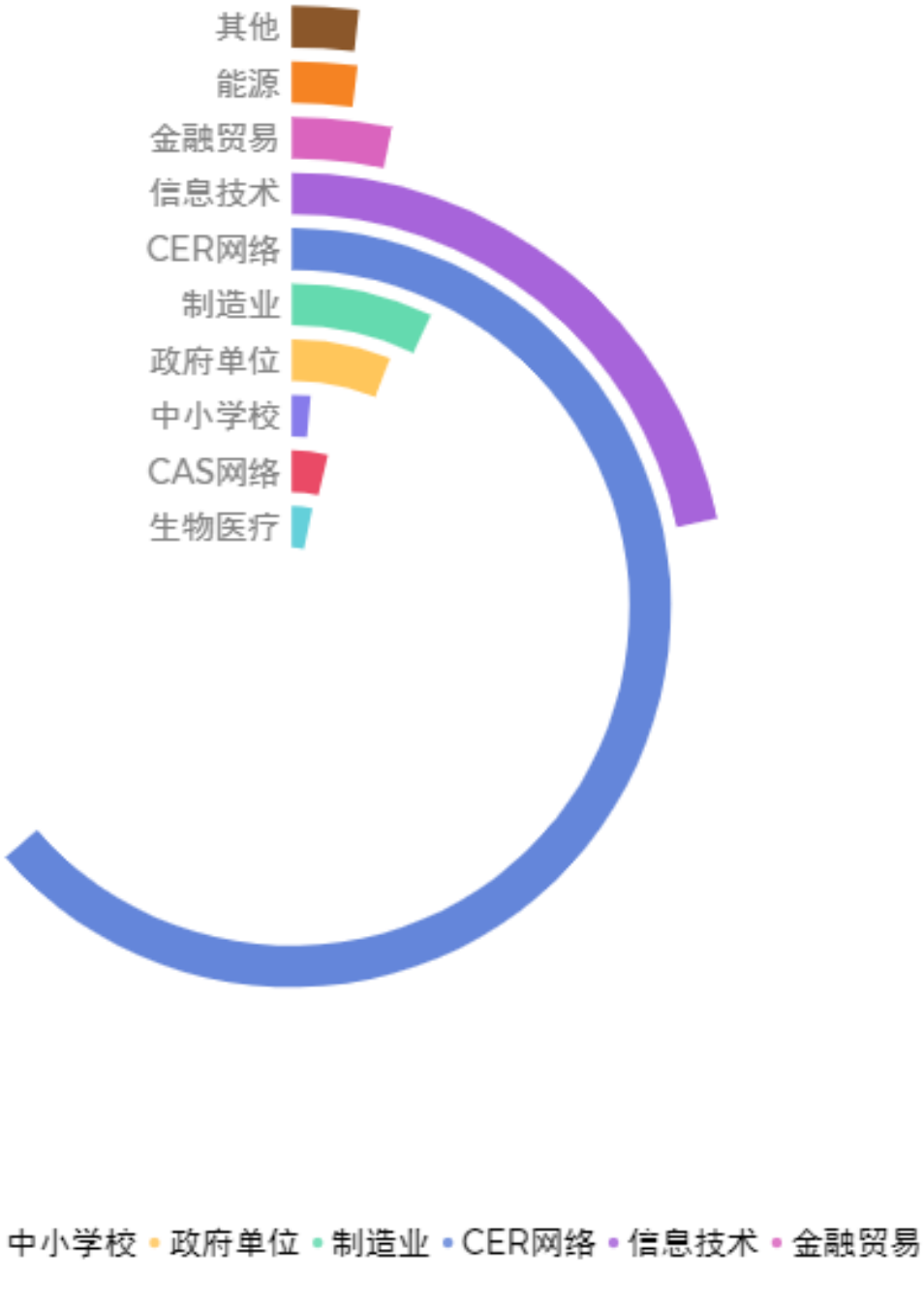
In summary, this sophisticated espionage script brings significant benefits to the attackers. After obtaining git credentials, they directly pull source code from the internal code server and package it together with scientific research data from the server, transmitting it to the springboard server.

xmrig

No attacker can resist installing a mining component on a Linux server equipped with five RTX4090 or eight RTX3090 graphics cards. Although our research team is not filled with geopolitical nationalists, this behavior makes us ponder whether the attackers intend to hinder the development of science and technology in our country. From the results, the presence of the xmrig component effectively conceals the true purpose of UTG-Q-008. After all, the various components mentioned above typically remain on the compromised machines for a maximum of five minutes during the actual attack process, while the xmrig component continues to run until we arrive to collect digital forensic evidence.

Scope of Impact

Based on QiAnXin's telemetry data, the number of affected IP addresses (specifically identifiable units) in the past three years has reached over 1500+. The highest proportion belongs to the Education Network (CER), aligning with the content of the UTG-Q-008 attack list.



We have also monitored some overseas affected IP addresses. However, due to the lack of sufficient localization methods, we can only identify overseas important units through IP reverse lookup:

单位
印度[redacted]研究所
印度尼西亚[redacted]大学
越南[redacted]公司
越南[redacted]公司
韩国[redacted]外科医院
越南[redacted]研究所[redacted]
巴西[redacted]息技术
美国[redacted]信息技术
厄瓜多尔[redacted]学院
伊朗[redacted]公司
印度尼西亚[redacted]公司
韩国[redacted]公司
奥地利[redacted]平台
捷克[redacted]网站
菲律宾[redacted]组织
印度教育承包商[redacted]

Overseas affected industries include educational contractors, universities, United Nations organizations, research institutes, information technology companies and so on.

Attribution

UTG-Q-008 follows standard working hours. In the UTC+8 time zone, attackers generally work from 14:00 to 19:00. However, overtime situations frequently occur during late nights, predominantly between 22:00 and 04:00 the next morning. It is speculated that the attackers are located in Eastern Europe. We are more inclined to believe that the botnet is of an outsourced or cooperative nature, while the actual "client" with a demand for scientific research data and source code remains hidden in the shadow.

During the expansion of the infrastructure, we discovered on a third-party platform that some payload from the Nishang framework had connected to some same springboard IP.

README License

Nishang

Nishang is a framework and collection of scripts and payloads which enables usage of PowerShell for offensive security, penetration testing and red teaming. Nishang is useful during all phases of penetration testing.

By Nikhil Mittal Founder of Altered Security - Hands-on red team and enterprise security training!

Usage

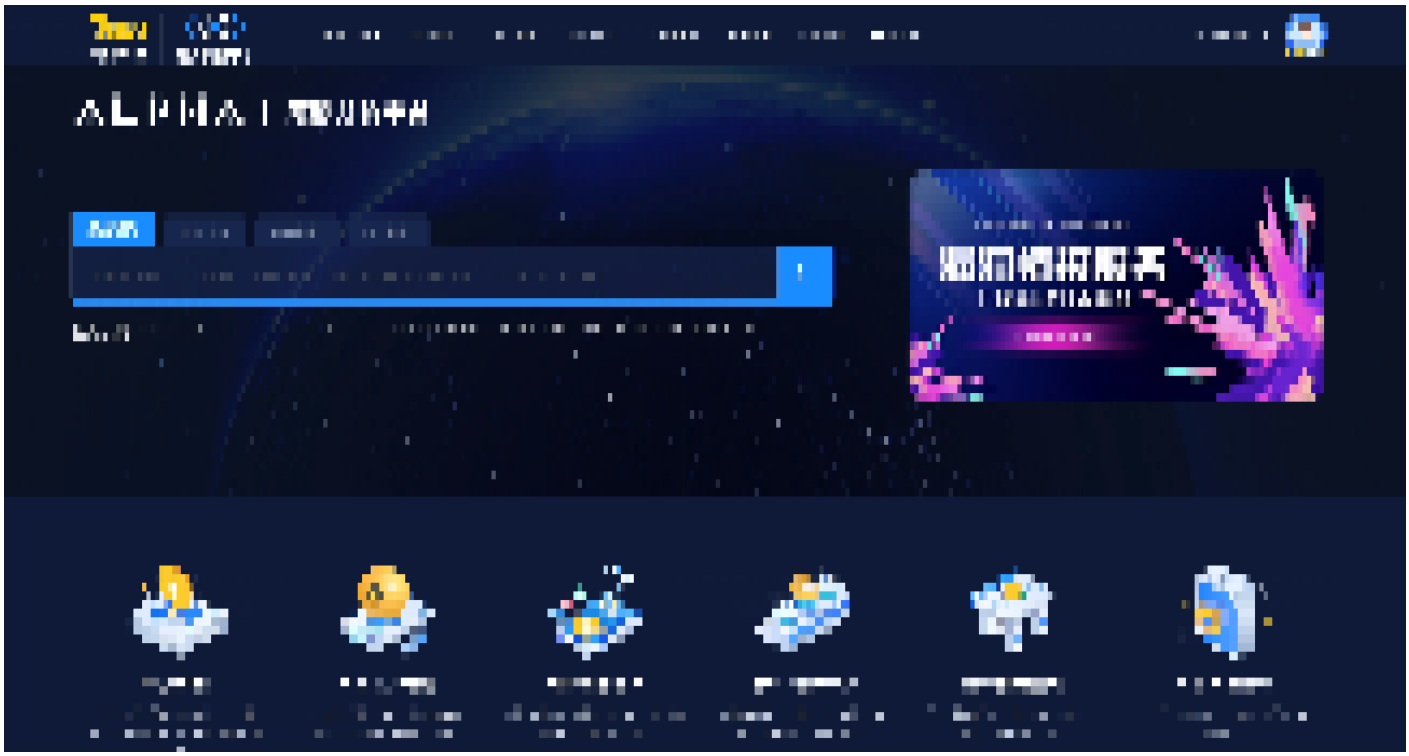
Import all the scripts in the current PowerShell session (PowerShell v3 onwards).

```
PS C:\nishang> Import-Module .\nishang.psm1
```

Reviewing the internal reports of the past three years, we found that the usage of this framework in espionage activities targeting domestic entities is very rare. Only APT-Q-78 employed the Nishang framework in 0-day attacks targeting China's scientific research field during the period of 2022-2023. This dual coincidence of target industry and weapon is intriguing, but it is not sufficient to be considered as highly reliable attribution.

Summary

Currently, all products based on QiAnXin Threat Intelligence Center's threat intelligence data, including QiAnXin Threat Intelligence Platform (TIP), TianQing, TianYan Advanced Threat Detection System, QiAnXin NGSOC, and QiAnXin Situation Awareness, fully support accurate detection of such attacks.



IOC

For commercial reports and victims related to UTG-Q-008, please contact QiAnXin Threat Intelligence Center (ti.qianxin.com).

Reference link

[1].https://ti.qianxin.com/blog/articles/The-Nightmare-of-EDR-Storm-0978-Utilizing-New-Kernel-Injection-Technique-Step-Bear-CN

[2].https://www.bleepingcomputer.com/news/security/fbi-disrupts-russian-moobot-botnet-infecting-ubiquiti-routers/

[3].https://www.sentinelone.com/press/sentinellabs-identifies-hidden-link-between-trickbot-anchor-purported-north-korea-lazarus-tool-deployment/