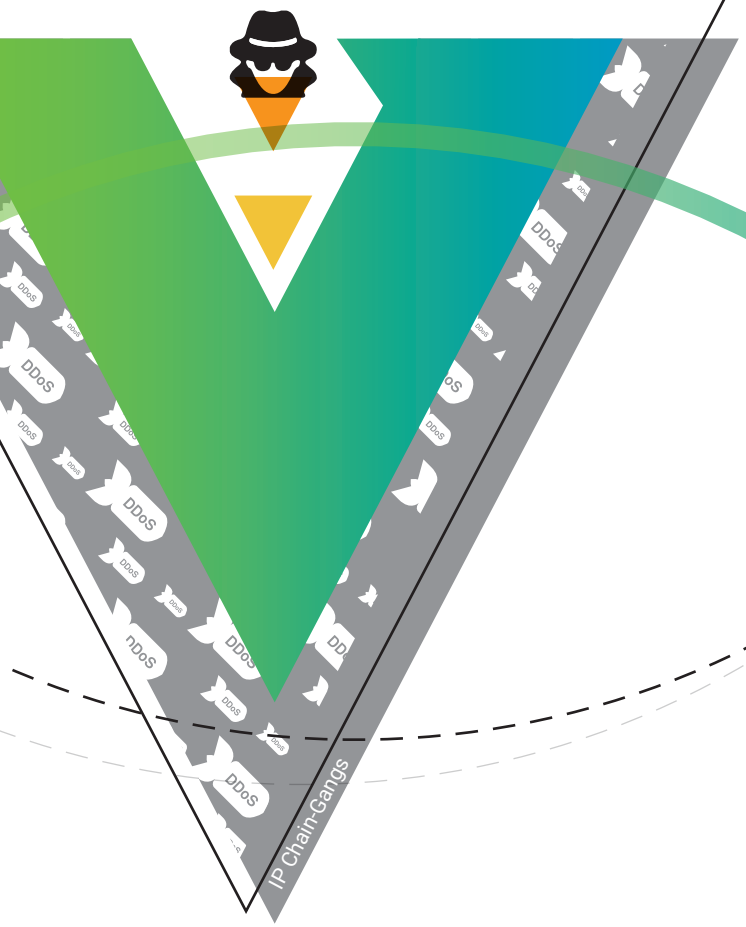


NSFOCUS

Behavior Analysis of IP Chain-Gangs

Hongbo Yang, Xiaobing Sun, Richard Zhao

NSFOCUS, Inc.

December 2018

NSFOCUS

About NSFOCUS

NSFOCUS, Inc., a global network and cyber security leader, protects enterprises and carriers from advanced cyber attacks. The company's Intelligent Hybrid Security strategy utilizes both cloud and on-premises security platforms, built on a foundation of real-time global threat intelligence, to provide multi-layered, unified and dynamic protection against advanced cyber attacks.

NSFOCUS works with Fortune Global 500 companies, including four of the world's five largest financial institutions, organizations in insurance, retail, healthcare, critical infrastructure industries as well as government agencies. NSFOCUS has technology and channel partners in more than 60 countries, a member of the Microsoft Active Protections Program (MAPP), and the Cloud Security Alliance (CSA).

A wholly owned subsidiary of NSFOCUS Information Technology Co. Ltd., the company has operations in the Americas, Europe, the Middle East and Asia Pacific.

Special Statement

All data for analysis is fully anonymized and no customer information is included and disclosed. For any potential concerns and issues, please feel free to contact the authors.

Behavior Analysis of IP Chain-Gangs

NSFOCUS

Hongbo Yang, Xiaobing Sun, Richard Zhao
NSFOCUS, Inc.
December 2018

Contents

1 Introduction and Executive Summary

In the *NSFOCUS 2018 H1 Cybersecurity Insights*[1] report, we observed that "Recidivists are responsible for 40% of the attacks, most of which are botnet activities and DDoS attacks". Since botnet activities and DDoS attacks are usually launched from multiple sources in a collaborative way, it's not surprising to see that many of these recidivists are working together as a group in these attacks. We call these groups "IP Chain-Gangs". Throughout this report we try to identify "IP Chain-Gangs" and then study their behaviors with the whole gang as the unit using DDoS attack data collected by NSFOCUS since 2017.

The logic behind this approach is that each IP Chain-Gang is presumably controlled by a single threat actor, or a group of related threat actors, and should therefore exhibit similar behaviors among the various attacks conducted by the same gang. By studying the historical behaviors of the gang, we hope to build a gang-profile that can help better describe how the threat actor(s) behind it operate, what their preferred attack methodologies and characteristics are, and how to build better defense against future attacks launched by them.

In this report, we introduce the IP Chain-Gang concept, and then focus on the statistical analysis of gang behaviors. From our analysis, we observed that:

- These gang members, though only a tiny fraction (2%) of all the attackers, are responsible for a much larger portion (20%) of all of the attacks.
- 20% of gangs are responsible for about 80% of all attacks launched by the gangs.
- Reflection attacks are the dominant attack methods favored by the gangs, specifically in high-volume attacks.
- Gangs typically do not operate at their full potential capacities. However, knowing their maximum attacking power is very important in planning the defense against them.

This report is the first in a series of the IP Chain-Gang topic. In future reports, we plan to examine how gang members have evolved and connected and how to apply that knowledge to build a more effective defense against them.

We believe that this is the first time that DDoS attacks are studied as coordinated gang-activities. As such, from this new view angle, we can gain a few unique insights on how DDoS attacks are conducted. In turn, this will help us better detect, mitigate, forensically analyze, and even predict DDoS attacks.

1 https://nsfocusglobal.com/2018-h1-cybersecurity-insights/

2 Identifying IP Chain-Gangs

To identify an IP Chain-Gang, we start by analyzing the DDoS attacks data collected by NSFOCUS since 2017 and undergo the following steps (for a more detailed discussion on this gang identification algorithm, please refer to the paper *Detection of IP Gangs: Strategically Organized Bots*[2]):

a. Identify the attackers who participated in one collaborated attack and place them into a group. Here we define a collaborated attack as those individual attacks that are against the same target at roughly the same time. Since these attackers are working together, it is plausible to believe that they are controlled by the same threat actor.

b. If two groups from the previous step overlap or their behaviors are significantly alike, we merge them into one bigger group. Repeat this group-merging process until no more significantly overlapping groups exist. A sophisticated machine learning algorithm is used to determine the threshold of the "significance".

c. Extract the core members of each attack group by purging "occasional attackers" (those attackers who participated in only a small percentage of attacks) from the group. These core members in an attacker-group become an IP Chain-Gang.

Through this process, we have identified over 80 active IP Chain-Gangs. In this study, we chose rather restrictive parameters in our algorithms. Therefore, all members in these gangs are really serious recidivists. Each of them has performed multiple attacks over our studying period. Consequently, though the number of these gang members is only about 2% of all the attackers in our dataset, they are responsible for about 20% of all the attacks.

It should be mentioned that the composition of any particular gang is very dynamic, as old members leave (presumably because the owner of the system removed the malware and patched the security hole used by the threat actor to gain access to the system originally) and new members join (new systems being infected by malware and become part of the botnet) over the time. In this report, we study the gang behaviors as if they are static entities over the study period. In future studies, we will take the dynamic nature into consideration.

2 https://www.researchgate.net/publication/326162077_Detection_of_IP_Gangs_Strategically_Organized_Bots

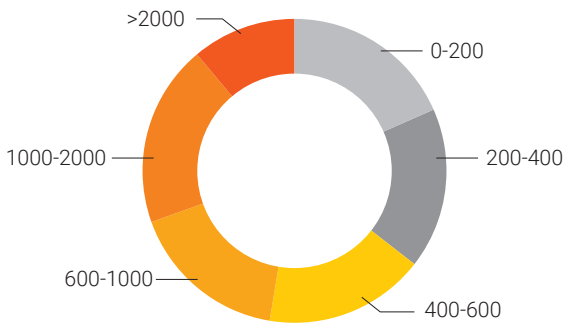
3 Statistical Analysis of IP Chain-Gangs

After the gangs are identified, we can study per-gang level behavior from several different perspectives. Unless otherwise stated, the numbers reported in this section are cumulative across all the members of the same gang.

3.1 IP Chain-Gang Size (Member Count)

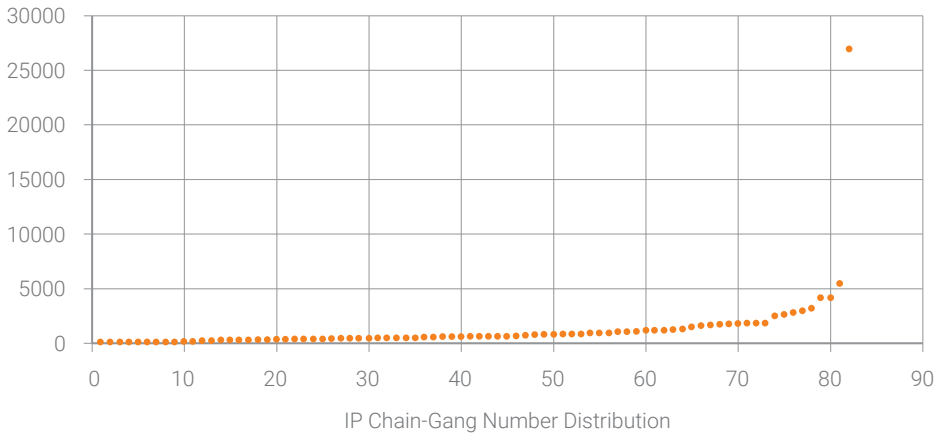
The chart below shows the size distribution of the IP Chain-Gangs. Most of the gangs have less than 1000 members, but we also see one gang with more than 26,000 members.

Figure 1 IP Chain-Gang Size Distribution



The chart below shows the distribution of gang size of all the IP Chain-Gangs we identified, sorted by the gang size. One dot in the chart represents one gang, for a total of 82 gangs.

Figure 2 IP Chain-Gang Size (Per-gang Sorted)



3.2 Total Attack Volumes

The chart below shows the distribution of the total attacking traffic volumes generated by the gangs, accumulated over all the attacks from the member of the same gang. While the attacking volume seems to vary drastically among different gangs, the majority of them generated more than 50TB total traffic over our study period.

Figure 3 Total Attack Volume Distribution

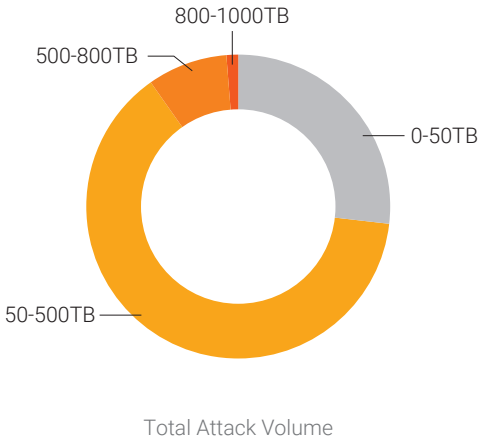
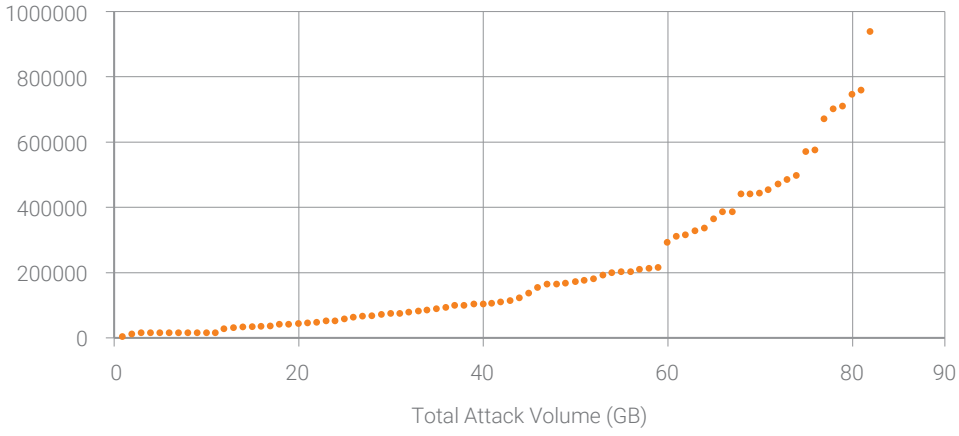


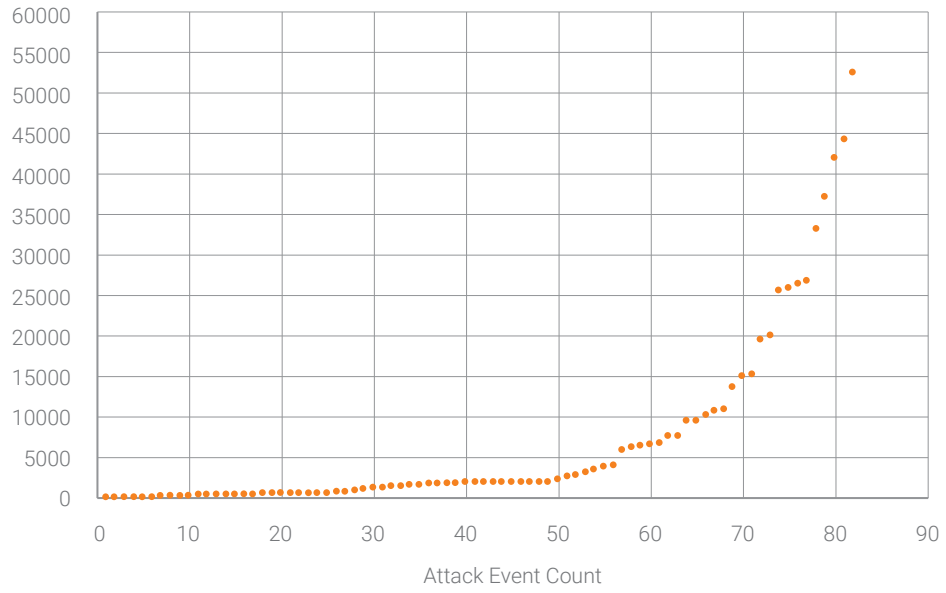
Figure 4 Total Attack Volume (Per-gang Sorted)



3.3 Total Attack Count

The chart below shows the number of DDoS attack events each gang launched, sorted by the event count. Unsurprised, roughly 20% of the gangs are responsible for 80% of the attacks.

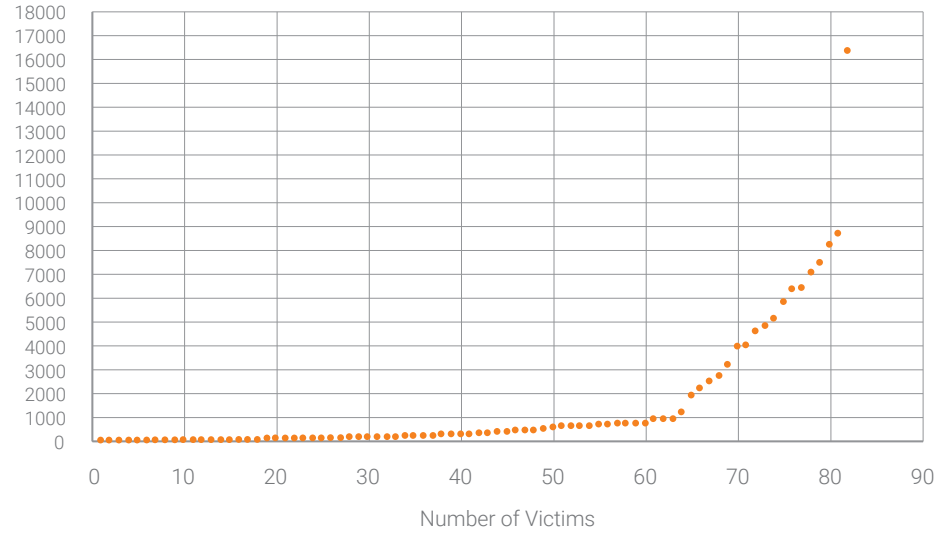
Figure 5 Total Attacks Count (Per-gang Sorted)



3.4 Number of Gang Victims

The chart below shows the number of victims each gang attacked, sorted by the victim count. We see 80% of gangs have less than 1000 victims, while one gang is responsible for attacking about 15% of the victims.

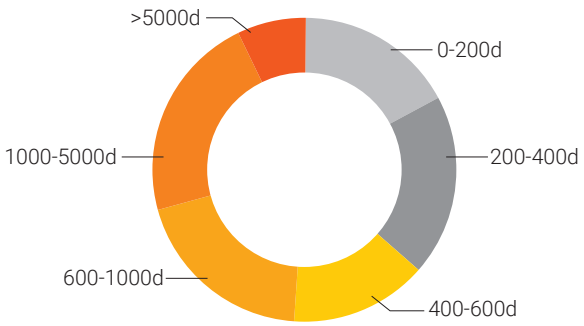
Figure 6 Number of Victims Count (Per-gang Sorted)



3.5 Total Attacking Time

The chart below illustrates the distribution of total accumulated attacking time from all the members of the same gang. While some gangs recorded total attack time of more than 5000 days, the majority of them recorded less than 1000 days.

Figure 7 Total Attacking Time Length



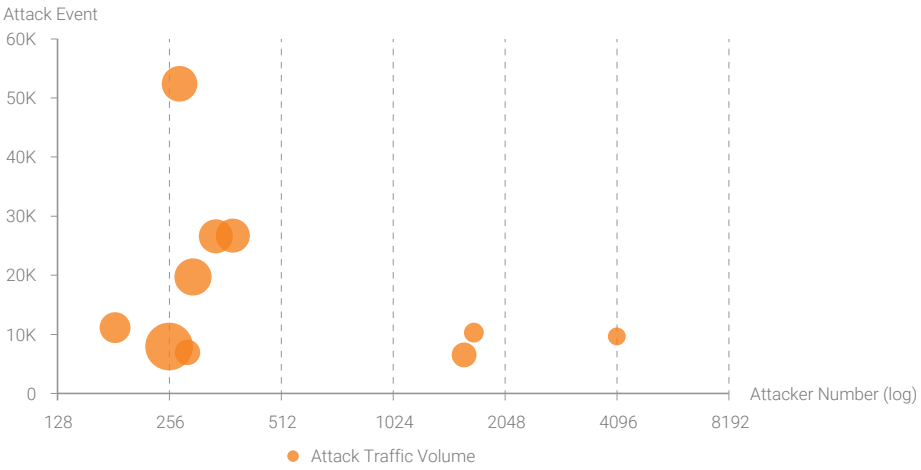
Gang Attacking Time

3.6 Comparing Gang Size, Attack Count and Attack Volume

While it is natural to assume that a bigger gang would launch more attacks with higher volume, that is actually not the case. As shown in the next chart, an IP Chain-Gang with less members may end up doing more attacks and sending more attacking traffic than that of a gang with more members.

Taking the top 10 gangs with the highest traffic volume, the chart below shows the IP Chain-Gang size on X-axes (log scale), the attack count on the Y-axes, and attack volume as the size of the orange bubble.

Figure 8 Gang Size vs. Attack Count vs. Attack Volume



Top IP Gang Attack Event Number - Attacker Number - Attacker Volume

As shown in the chart, the bigger the bubble size does not correspond to more attacker number or attack count. From the chart, we can see a gang with 274 members attacking at an extremely high frequency (> 50K) which exceeds all others, while the biggest bubble (i.e. the biggest attack volume) has less members (256) and less attack count (< 10K). This suggests that the attackers in this particular gang probably have bigger pipes at their disposal.

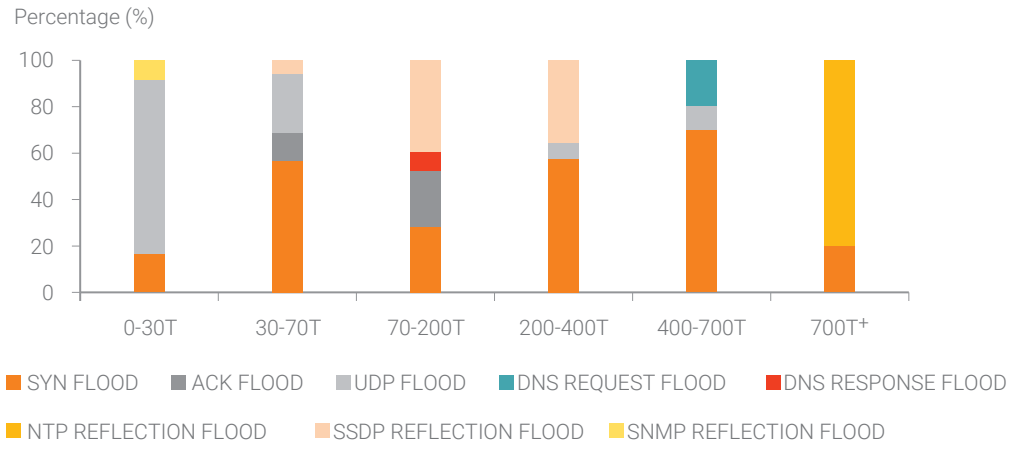
3.7 Attack Types (Methods)

The attacking method(s) used is another important aspect when studying the DDoS attacks. Different methods have different characteristics such as traffic volume generated, ease of implementation and detection, system dependencies, etc.

3.7.1 Attack Type vs. Attack Volume

The next chart shows the composition of various attack types across different attack volume ranges.

Figure 9 Attack Type vs. Attack Volume



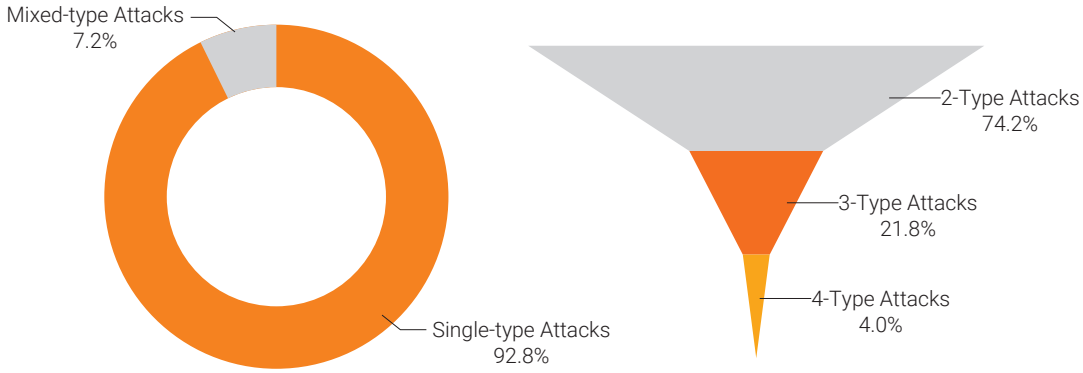
IP Gang Attack-Type Classification against Attack Volume Size

NTP reflection flood type attacks make up most of volume in the high-volume attacks (due to their great amplification factor), while SYN flood attacks spread across all the spectrums (most probably due to their simplicity). These two attack types along with UDP flood and SSDP reflection flood make up the majority of the attack types across the board.

3.7.2 Single-type Attacks vs. Mixed-type Attacks

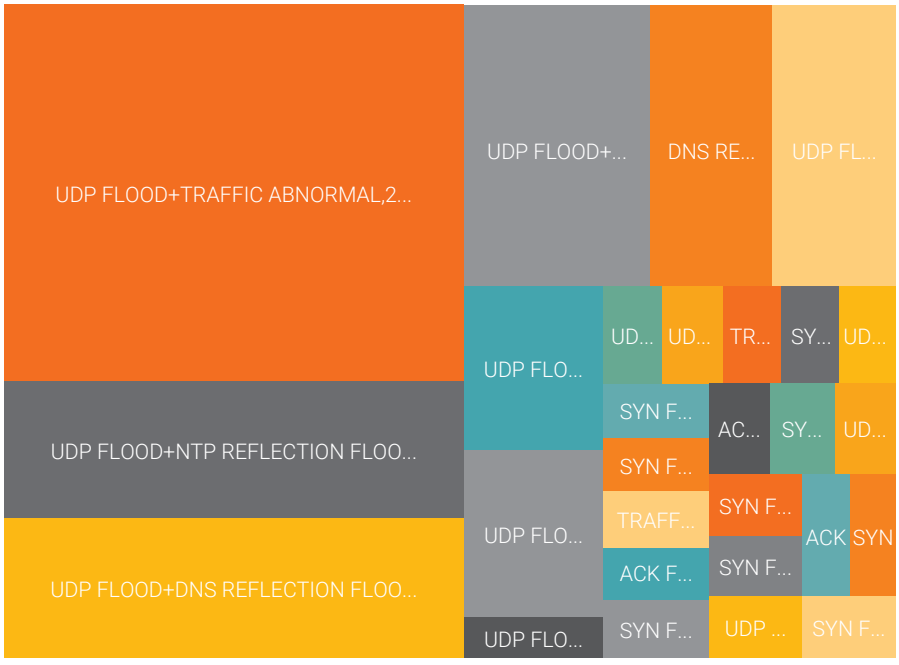
Many gangs have their preferred attacking method. This obviously reflects the skills and preference of the threat actors behind the gang. However, we also see some gangs employ multiple attacking methods in their attacks, sometimes even mixing the attack methods in one attack. The two charts below show the attacking methods used by one specific gang.

Figure 10 Single-type Attacks vs. Mixed-type Attacks (One Gang)



Single-type Attacks vs. Mixed-type Attacks

Figure 11 Combinations of Different Mixed Types



Common Combinations for Mixed-type Attacks

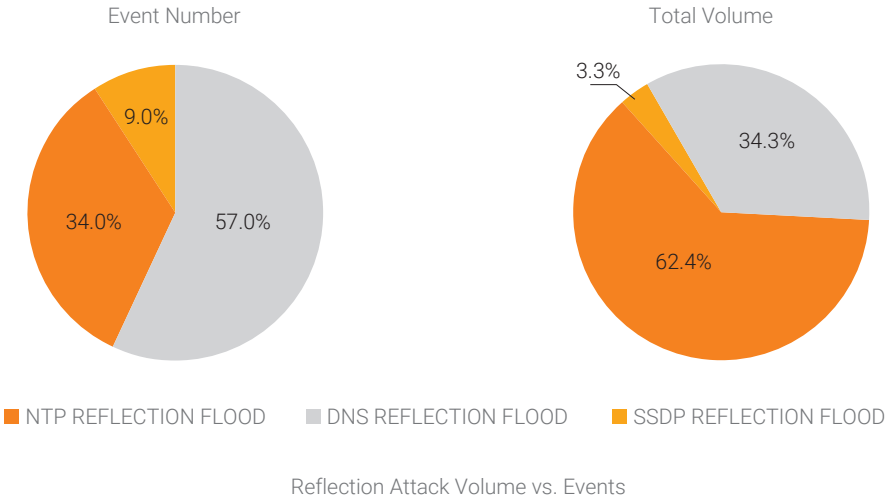
In the first chart, we observed that most (93%) of the attacks launched by this gang use only a single attack method. However, every now and then multiple attack methods are used together in one coordinated attack against a single victim. Of all the mixed attacks, ¾ of them involve two different attack methods, and 4% of them involve 4 methods.

In the second chart that shows the different composition of mixed attack methods, we see UDP flood spread throughout. This is not a surprise, as UDP flood is quite a long-lived DDoS attack method and still enjoying popularity today.

3.7.3 Reflection Attack Volume vs. Events

We see that various types of reflection attacks are being used more and more in DDoS attacks, especially in high volume attacks. Moreover, we see some gangs use multiple reflection attack methods together. The chart below is one example:

Figure 12 Reflection Attack Volume vs. Count (One Gang)



Reflection Attack Volume vs. Events

From the attack event count perspective, DNS reflection flood claims 57% of overall reflection attacks followed by NTP reflection flood. In terms of volume, the NTP reflection flood takes the lead at 62.5%, DNS is a close 2nd. NTP reflection flood is the stronger DDoS attack, given its capability to trigger higher traffic volume.

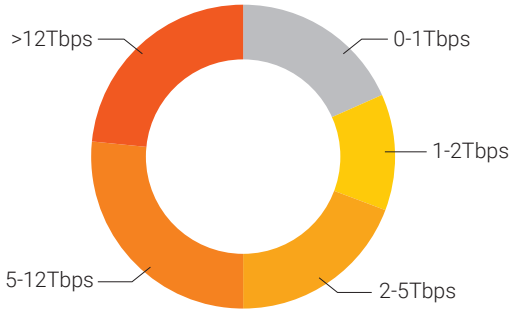
3.8 The Peak Rate

The peak rate of the attacking traffic generated by a gang is another important aspect to study, as it indicates the gang's ultimate attacking power against a target. In the following subsections, we look at the peak rate distribution of all the gangs, as well as how the peak rate changes over time.

3.8.1 All Gangs' Peak Rate

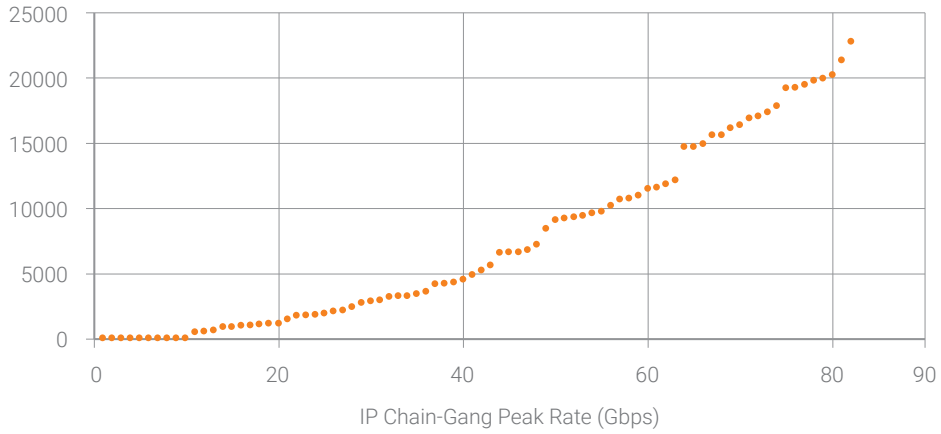
The peak rate (in Tbps) is the key parameter indicating the strength and viciousness of a certain group. Most IP Gangs have peak rate over 2 Tbps.

Figure 13 IP Chain-Gang Peak Rate Distribution



IP Chain-Gang Peak Rate Distribution

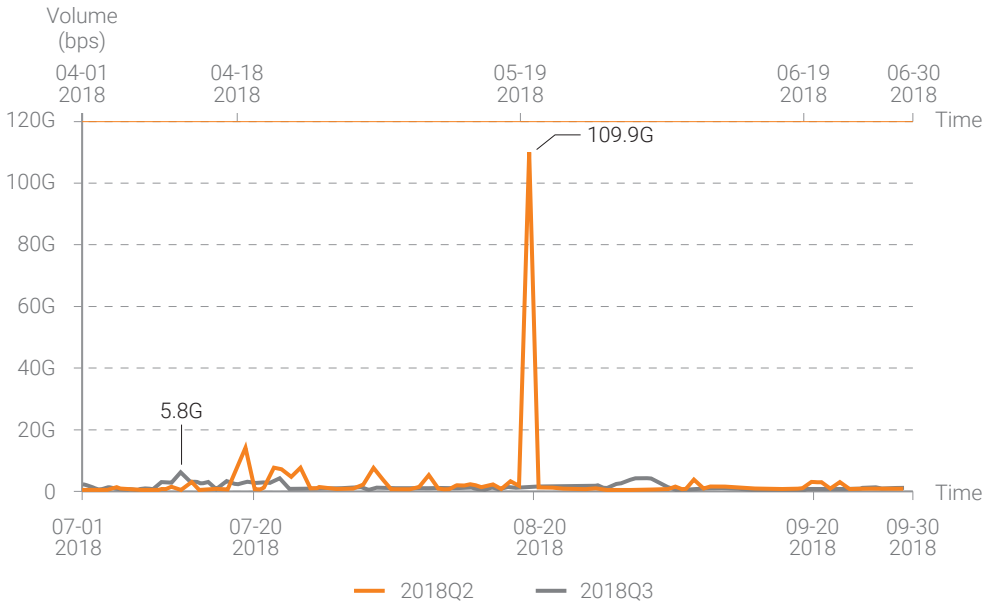
Figure 14 IP Chain-Gang Peak Rate (Per-gang Sorted)



3.8.2 Single Gang Attack Peak Rate

It should be noted that not all attacks launched by one gang operate at their potential peak rate. In fact, most attacks are at a much lower level, due to both the threat actor's "business need" and the available members of the gang's at the time. For example, the following chart compares one particular gang's peak attack rate in two quarters:

Figure 15 Single Attack Peak Rate Trend (One-gang)

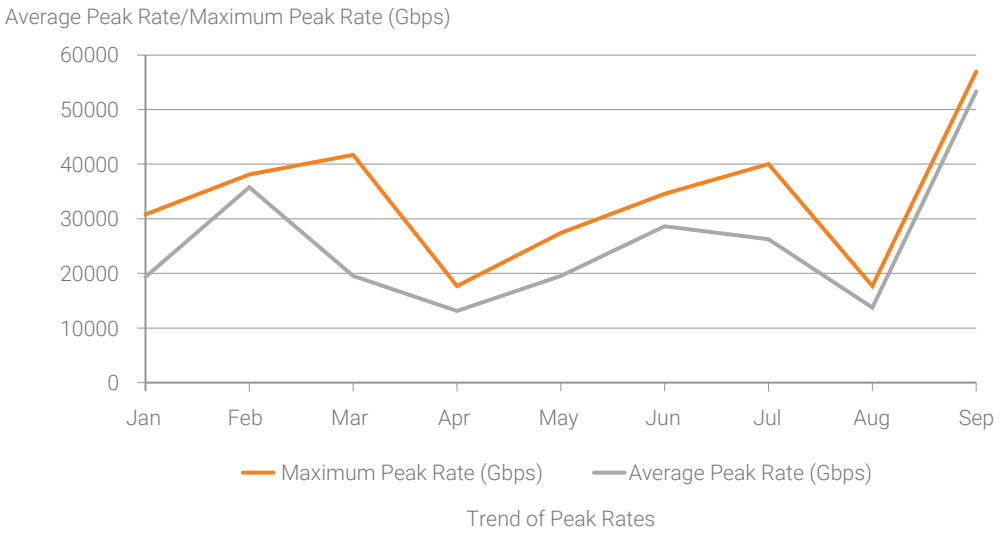


This particular gang had its peak attack rate on May 20th, 2018. It has much lower activities in other months. Although there is no clear trend on the peak rate for this gang, the difference between its usual peak rate and its top peak rate is colossal. In other words, when this gang is fully engaged, it's going to be quite powerful.

3.8.3 Top Ten Gangs

The next chart looks at the top ten gangs' peak attack rates during the months of January through September, 2018. With peak rates accumulated every day, the average and maximum peak rates over these months are then plotted.

Figure 16 Top Ten IP Chain-Gang Peak Rates



Maximum peak rates and average peak rates represent how strong and durable the IP Chain-Gang is. With the changing numbers through the past months in 2018, it rises and dips during the whole period which indicates the level of activities of these gangs.

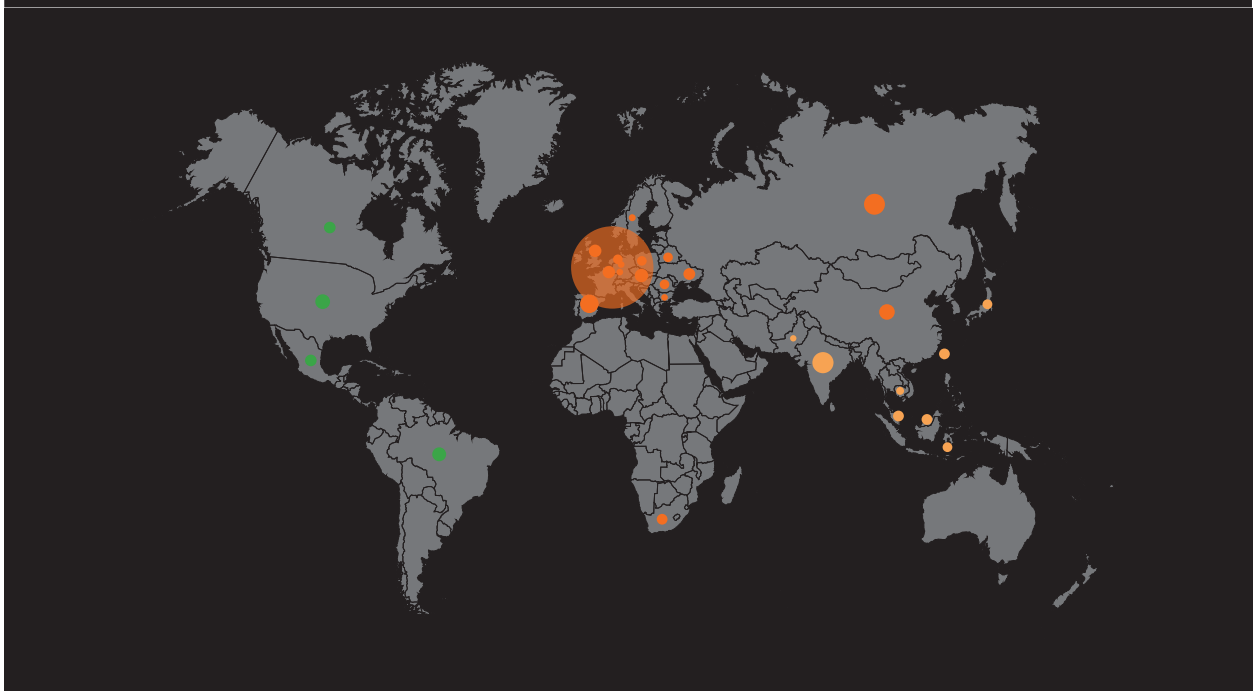
3.9 The Geo-Locations of Attackers and Victims (Excluding China)

For both attackers and victims, the geo-location information reveals where the valuable targets are and where the attacks "initiate". It may not be where the controlling threat actor(s) are located, but at least it shows the hot-spots where the DDoS activities occurred.

Since most of our sensors are located in China, it is not surprising that in our overall data, most of the attack sources and destinations are within China. To show the non-China activities, we took those data generated from sensors located outside China and study where they are located.

3.9.1 Distribution of Attacking Countries

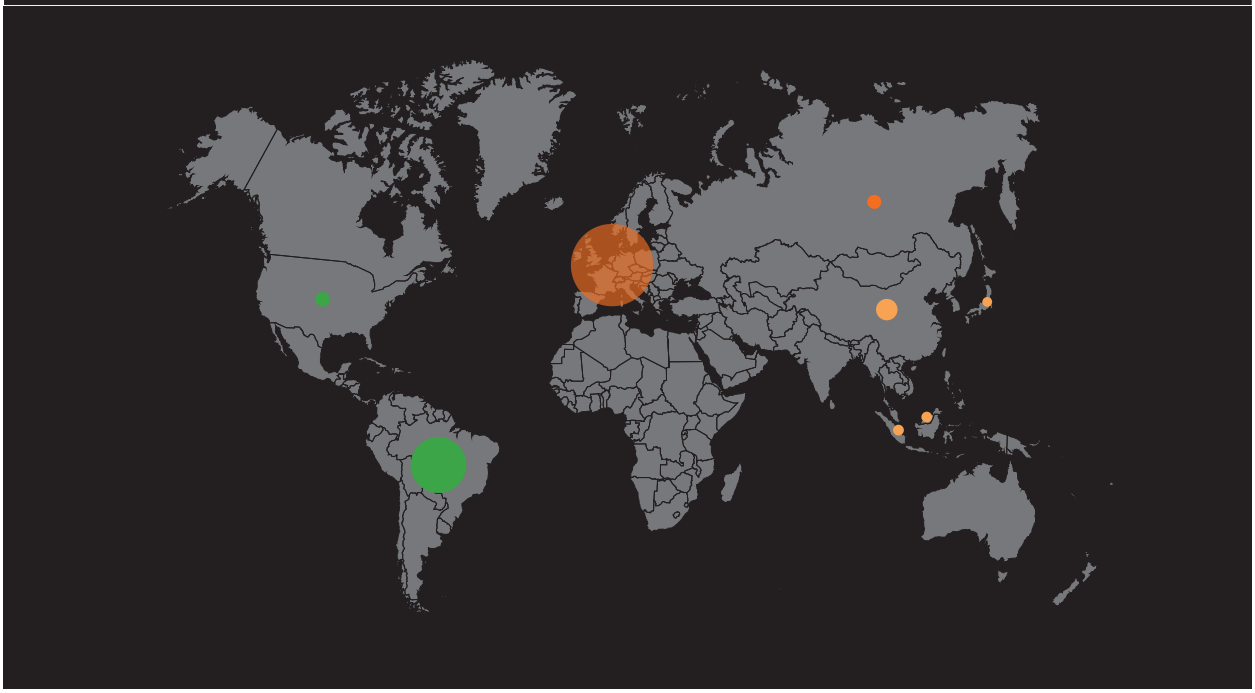
Figure 17 Distribution of Source Countries



From the attacker point of view, the No. 1 attacker source region are European countries. Asian countries as well as countries in North America also contributed a significant amount.

3.9.2 Distribution of Victim Countries

Figure 18 Distribution of Target Countries

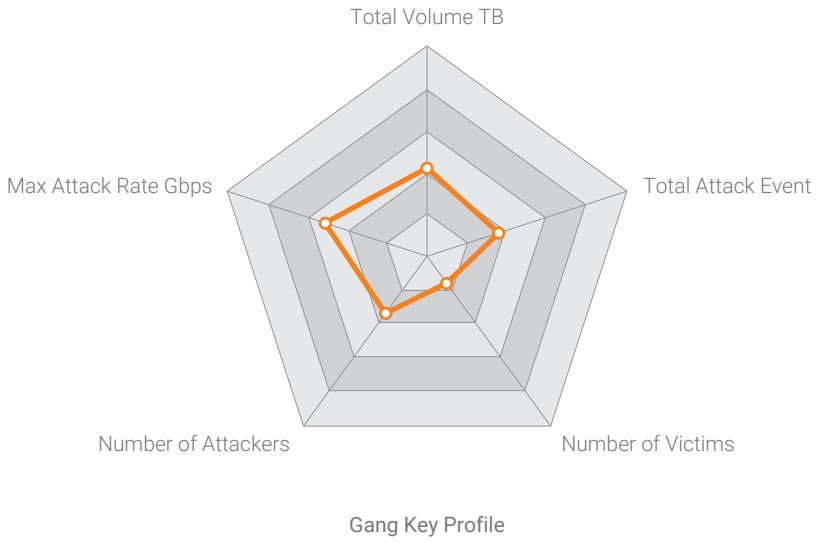


The attack victim distribution map shows Europe suffers the majority of the attacks while South America follows. There are also attacks to Asian countries as well as to the U.S..

4 IP Chain-Gang Profile Model

Taking the five (5) quantifiable features discussed above, we can make a radar chart like this: the value at the edge of this radar chart is the maximum value among all gangs for that specific feature. The center represents the 0 value of all the features. Each step on the axes is 20% of the maximum value. Then plotting a specific gang's relative feature value on the radar and connecting the dots, we get a full "profile" for that gang, as shown below.

Figure 19 IP Chain-Gang Profile Model



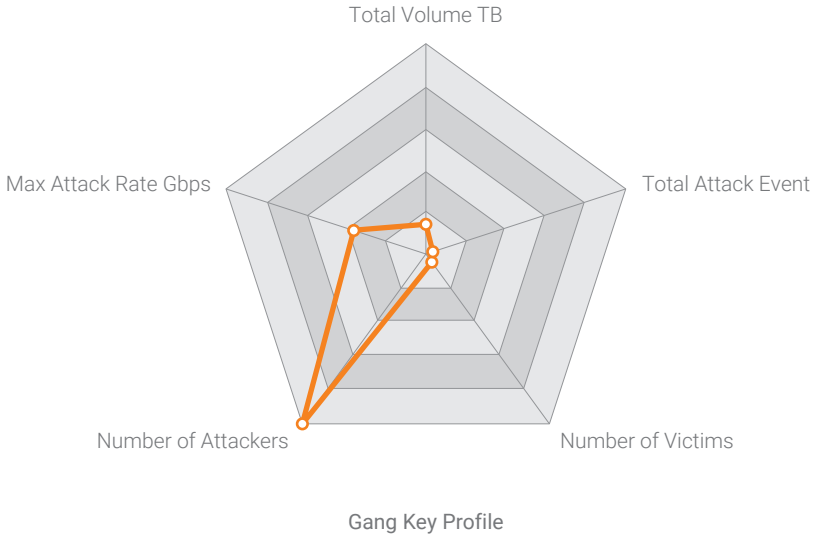
Now we can see different characteristics of a specific IP Chain-Gang in a multi-dimensional way to include attacking volume, number of events, victim number, attacker IP number and max attacking rate. A larger area surrounded by the orange lines normally indicates that gang is more intensively involved in the attack. It outlines the general overview on how big/strong the gang is. For example, the gang profiled in the chart above has a smaller number of attackers and victims but generated a larger number of attack volumes and has an overall higher peak rate, all relatively speaking.

In the rest of this section, we will take a look at three gangs, each with an extreme characteristic within its profile.

4.1 The Largest Gang

The chart below shows the profile of the largest gang we identified. From the chart, we can see that it didn't attack many victims, or launch many attacks. But it did have a higher peak attack rate, perhaps due to its large member base.

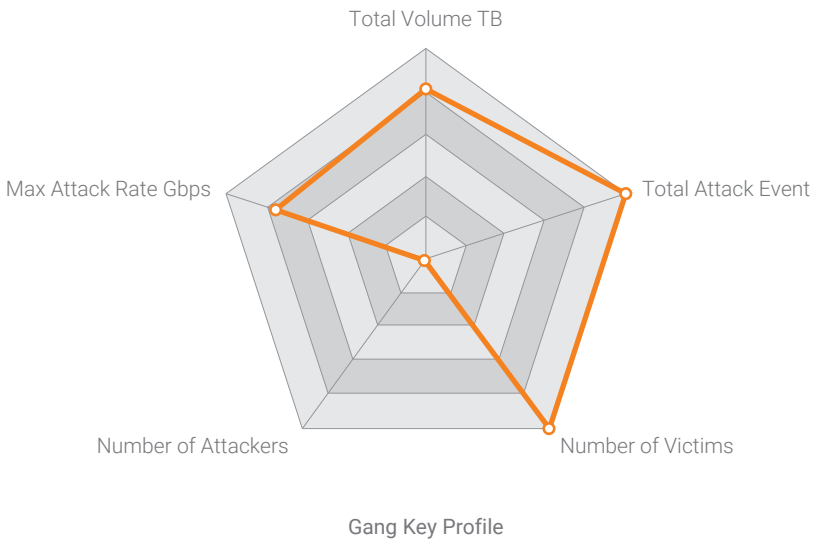
Figure 20 IP Chain-Gang Profile Model (Largest Gang)



4.2 The Most Active Gang

The gang with the greatest number of attacks and the greatest number of victims looks like below. It is actually a relatively small gang, yet it is capable of generating very large attacking volume and a high peak rate. This is certainly a very fierce gang.

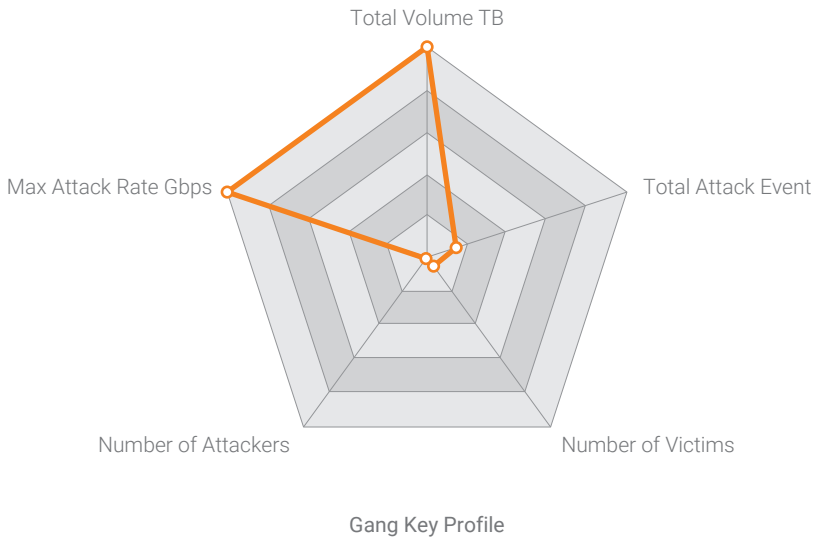
Figure 21 IP Chain-Gang Profile Model (Most Active Gang)



4.3 The Gang with Largest Volume

The gang below is responsible for the most attack volume and highest peak rate, despite its relatively small size and small number of victims and attacks. We can imagine that members in this gang probably all have access to a high bandwidth pipe.

Figure 22 IP Chain-Gang Profile Model (Largest Volume)



5 The Future Work

In this report, we presented a new view angel of analyzing cyber-attack activities based on the IP Chain-Gang concept. We attempted to show how a group of attackers controlled by a threat actor can operate and attack in a coordinated manner. We also presented a profiling model to help describe and compare these gangs.

Moving forward, we plan to track IP Chain-Gangs' evolving history and study the inner-connections among their members. It will help predict the attacks from those gangs, as well as build stabilized knowledge to make effective defense against them.

6 References and Acknowledgements

6.1 References

1. NSFOCUS 2018 H1 Cybersecurity Insights, by NSFOCUS Security Lab, https://nsfocusglobal.com/2018-h1-cybersecurity-insights/
2. Detection of IP Gangs: Strategically Organized Bots, by Tianyue Zhao, Xiaofeng Qiu. https://www.researchgate.net/publication/326162077_Detection_of_IP_Gangs_Strategically_Organized_Bots

6.2 Acknowledgements

We would like to thank the following colleagues for their invaluable assistants in preparing and analyzing the data, as well as in writing and reviewing this report: Lin Xu, Xue Mei, Cai Huayu, Alex Wang, Guy Rosefelt, Adeline Zhang.

For questions and comments, the authors can be reached at Xiaobing.sun@nsfocusglobal.com.

NSFOCUS

Security Made
Smart and Simple

www.nsfocusglobal.com