Metric design for botnet mitigation

Santiago Aragón*, Daniel Meinsma*, Pallavii Jagannatha* Owais Ahmed*,
* University of Twente, Nederlands

***Abstract*—**

***Keywords*—IEEEtran, journal, LaTeX, paper, template.**

I. What security issue does the data speak to?

Companies and people have more and more problems defending against cyber incidents. It is not possible to defend against all possible cyber attacks, so people have to make strategic decisions on where to invest in their security. Making strategic decisions about security requires us to have a clear understanding of these security issues and for this there is a great need for security metrics [1].

To be able to make metrics, we need reliable data. One way getting reliable data about cyber threats is capturing cyber attacks with a honeypot. Honeypots are systems made vulnerable intentionally to attract attackers to get inside the system to track their activities [2].

This paper proposes meaningful metrics based on the logs from the honey pot called *Elastichoney*. Elastichoney is a elasticsearch honeypot containing the RCE Remote code execution vulnerability -CVE- 2015-1427, which allowed attackers to execute java based code in the search bar of the honeypot [3]. The log files of this honeypot contains data collected over two months and tracked about 8k attempts to attack from over 300 unique IP addresses [1].

One of the major security issues that the digital world is currently facing is BOTNETS. In the paper *Measuring the cost of cybercrime* you can read that:

Botnets provide a versatile platform for a variety of criminal business models, including sending spam, committing click fraud, harvesting account credentials, launching denial-of-service attacks, installing scareware and phishing [4].

The authors of this paper think that the logs of Elastichoney tracks the activities of one or more botnets, because the following two features can be found in the attack patterns in the logs: First, by using the code execution vulnerability, the same operation steps are performed on multiple different files:

- Downloading the file using the command wget,
- Making the file location accessible by using the command chmod 777,
- Executing the file,
- Removing the file using the command rm.

Second, the operations on files were requested by multiple different sources.

These two features hints at distributed automated attacks, which is a clear indicator of botnet activity.

II. What would be the ideal metrics for security decision makers?

The security issue defined in section I have distinct effects depending on the issued party. We identify several parties that might be aimed by the security issue and we arrange them in two groups which are affected in similar ways. For each group, we identify how the botnet (BN) is negatively influencing their interests, as well as a set of ideal metrics which could help each party to derive better decisions to take the more effective counter measures to mitigate their problem while maximizing the return on security investment.

a) Internet Service Provider (ISP) & Botnet's victim: For this parties it is of main interest to avoid misuse of resources i.e. the ISP would like to block not legitim incoming traffic to its network while the botnet's victim would like to maintain availability of its resources.

b) Metrics:

- Attack prediction.
- BN identity detection.
- Resources (un)availability cost.

c) Regulator authority & Law enforcement agencies: The parties interested in prevent, prosecute and punish the ones behind a security issue, i.e., For the regulator authority is of main interest to transfer the risk to the ISP, to encourage the zombies machines to increase their security protection, while for a law enforcement agency is to have mechanisms to hunt down the attacker

d) Metrics:

- BN location detection

III. What are the metrics that exist in practice?

In the previous section we have seen which kind of metrics would be ideal for different parties facing the botnet issue. Before going to the metrics proposed for each party by this paper, first a list is presented of metrics that are already used by honeypots for botnet detection:

e) Network Fingerprinting: With this methods, you can create a metric that states which hosts communicates to which hosts. For example, you can track all the IP addresses , the honeypot communicates to . This is interesting because using this method , you can differentiate between different types of botnets for example if the honeypot is part of a traditional command and control botnet. The honeypots will only communicate to the controller and in a peer to peer botnet , the honeypots will create multiple other members of the botnet [5].

f) IRC related features: With this method, you can create metric that differentiates between a member of a member of a IR C type botnet and a non-infected member because these type of botnets send and receive signature commands over IRC channels [6].

g) Longitudinal tracking: With this method, you can create metric where you can visualise the number of attacks originating from a particular geographical location [6].

h) DNS tracking: This method is almost the same as Longitudinal tracking, but instead of tracking the geographical location, this method tracks the domain names.[AM2006]

i) Port Scan tracking from IDS logs: With IDS logs, you can view how many times a host tries to communicate with the honeypot and deduce if a port scan is active and create a metric that state show many port scans each communicating host performs [2].

j) Botnet resources tracking: In figure 1,you can see different resource aspects of botnet and each of these resources can be used to create metrics which differentiates between different types of botnets [5].

k) Botnet infection vectors: This method is almost the same as the botnet resources metrics but in this metrics, you differentiate between infection vectors of different types of vectors. See figure 2 and an extra example might be phishing [5].

l) Signature tracking: In a few methods seen above, we have seen metrics that could differentiate between different types of botnets and if we use a signature of a specific botnet, for example the backdoor port of a trojan horse that it uses , we can create a metric that tracks the infection rate of that specific botnet [7].

m)DNS sinkhole: If a honeypot gets infected by a traditional command and control botnet and uses a DNS server to communicate with this honeypot ,one can try to change the DNS location of the botnet controller to DNS location of the honeypot. This results in all the other members of the botnet in the same DNS server communicating to the honeypot instead of the botnet controller which results in a precise mapping of all the members of the botnet in the DNS server [2].

IV. A definition of the metrics you can design from the dataset

Our main focus for the analytics will include identifying clear statistics on events from whom, from where and what attacks recorded in the JSON logs available. The following metric definitions will help us critically analyze the nature and characteristics of attacker behavior, the means of predicting the attack, the severity of the incident and resources needed to mitigate the impact of this vulnerability.

n) Attacker Behavior:: The inputs for this metric will include geographical locations / regions such as continent, country and city name. The different methods and types of various user agents used to exploit the vulnerability paying particular attentions to the application names. The majority of source IPs sending data traffic. The form and payload strings matching the actual exploit. The number of recurring attacks from particular IPs. The most important of all is the breakdown of the individual payload and its various instruction sets. Thus by analyzing these trends and rankings of different key data fields, one can predict the methodology and characteristic of the attackers behavior.

o) Attack Prediction / Probability:: The required data for this metric will include specific port numbers that are being used including the majority of source IPs sending data traffic via URLs with payloads attached. The method and nature of data transfer via URLs for direct path to vulnerabilities that are exploitable. The number of port scan attempts without including payloads and recurring attacks from particular IPs. These datasets will help to derive patterns for attack prediction and probability of successfully exploiting the vulnerability.

p) Scale of Incident:: The inputs for this metric will include the payload strings matching the actual exploit. The amount of valid and invalid vulnerabilities, malicious IPs and the percentage of false positives. The sum of individual attacks within a specific time interval. This metric will therefore help to determine the nature and level of attack strength and effort put in to gain access to critical infrastructure.

V. An evaluation of the the metrics you have defined. This should include graphical representations of the metrics (e.g., histograms, scatter plots, time series, bar charts).

VI. Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

References

[1] R. Böhme, "Security metrics and security investment models," in *Advances in Information and Computer Security*. Springer Berlin Heidelberg, 2010, pp. 10–24.

[2] P. Wang and et al., "Honeypot detection in advanced botnet attacks," in *nternational Journal of Information and Computer Security 4.1*. USENIX, 2010, pp. 30–51.

[3]

[4] R. Anderson and et al., "Measuring the cost of cybercrime," in *The economics of information security and privacy*. Springer Berlin Heidelberg, 2013, pp. 265–300.

[5] Grizzard, J. B., and et al., "Peer-to-peer botnets: Overview and case study," in *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, 2007.

[6] M. Abu-Rajab and et al., "A multifaceted approach to understanding the botnet phenomenon," in *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. ACM, 2006.

[7] M. Andreolini and et al., "Honeyspam: Honeypots fighting spam at the source," in *Proceedings of the Steps to Reducing Unwanted Traffic on the Internet on Steps to Reducing Unwanted Traffic on the Internet Workshop. USENIX Association*. USENIX, 2005.