Visualizing Amazon Web Services Honeypot Data

Menuka Herath, GRA; Gopi Chand Ganguru; Sanjana Taduri

College of Business + IT, MSIT Program

INTRODUCTION

The importance given to cybersecurity is evolving day by day[1]. Researches prove that the number and seriousness of cyberattacks are expected to grow over the years[2]. Cyber attacks has become a significant threat to governments and countries due to the increasing number of data breaches reported around the world. Consequently cybersecurity has become one of the widely spoken topics today. Countries have been encountering difficulties in preventing cyber attacks as it is now evolving rapidly across different areas as such, Business, Government, Military, Healthcare, Legal, Financial etc. The costs to resolve cyber attacks are equally high compared to the number of attacks reported.

The importance of cybersecurity can be studied closely with reference to following figures. As the data analytics forecast results imply, the number of attacks are expected to grow over the following years.

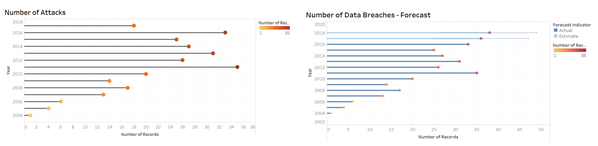

Fig. 1 Source : Kaggle – Data Breaches

Following Figures explain most common data breaching methods and different sectors that data breaching have affected.

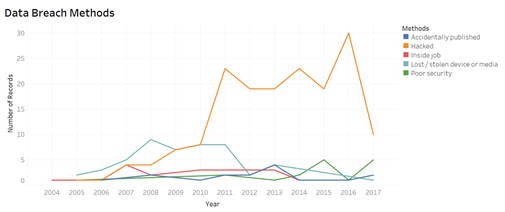

Fig. 2 Source : Kaggle - Data Breaches

Figure 2 presents the trend of data breach methods reported over the years. The sum number of incidents reported for each method is represented by a color. Hacking has been the reason for the highest number of incident reports over the years.

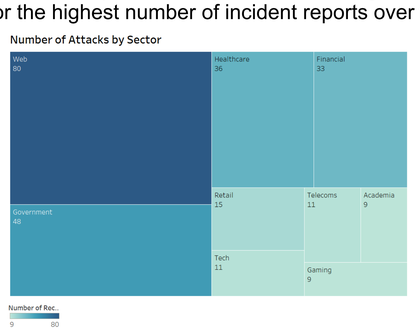

Fig. 3 Source : Kaggle - Data Breaches

Figure 3 is a visual display of the sum number of attacks by sector. Color and the size shows the sum number of Records.

METHODS

This project is conducted to gain better insight on cybersecurity and its' threats. Our project objective is to explore how effective "honeypot" is, as a security mechanism. Honeypot is used to secure computer systems from threats, attacks and prevent unauthorized user logins. Our curiosity comes from how it functions and its effects on cybersecurity.

We utilized Tableau software to analyze data, and present our project results.

The primary purpose of honeypots are to attract cyber threats, and it operate as a real computer network with less protection[3]. it is detached from the company network and closely monitored to identify emerging threats. It can prevent attackers accessing the internal network.

Honeypot Architecture

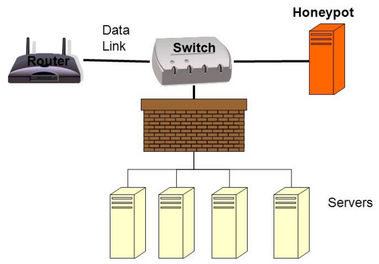

Fig. 4 Honeypot Architecture

Setting and context

The AWS Honeypot dataset for this project is acquired from Kaggle data source and the dataset includes 451,582 data points from 3rd March 2013 to 8th September 2013. The data is broken down to discover the research questions in identifying top threats based on IP addresses, frequency of attacks, mostly attacked destination ports, mostly used protocols and ports to perform the attack.

References

[1]Fischer, E. A. (2016, August 16). *Cybersecurity Issue and Challenges*. Retrieved from https://fas.org/sgp/crs/misc/R43831.pdf

[2]Hartwig, R. P., & Wilkinson, C. (2014, June). Retrieved from Cyber Risks: The Growing Threat: https://www.iii.org/sites/default/files/docs/pdf/paper_cyberrisk_2014.pdf

[3]Kartha, A. (2015, September 04). *Hunting hackers with honeypots.* Retrieved from ITProPortal: https://www.itproportal.com/2015/09/04/hunting-hackers-with-honeypots/

[4]Martin, G. (2014, September 15). *How honeypots can help stop data breaches* . Retrieved from Government Product News : http://americancityandcounty.com/security/how-honeypots-can-help-stop-data-breaches-related-video

[5]Martin, G. (2014, September 01). *How to Use "Honeypots" to Overcome Cybersecurity Shortcomings* . Retrieved from POWER: Business & Technology for the Global Generation Industry Since 1882: http://www.powermag.com/how-to-use-honeypots-to-overcome-cybersecurity-shortcomings/?pagenum=1

Analysis

Geographical Analysis

Mostly Attacked IP Addresses

Figure 5, Geographical Analysis indicates the results of the 10 mostly attacked IP Addresses. The countries are color coded based on the index function from 1-10.

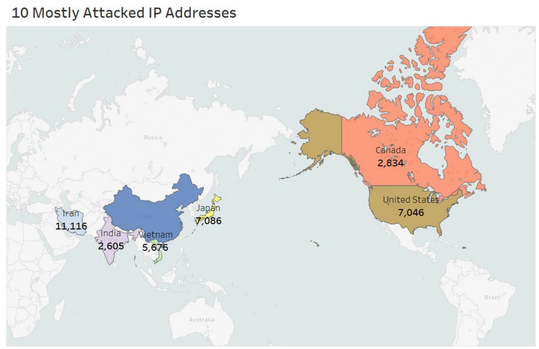

Fig. 5 Mostly Attacked IP Addresses

Timeline Analysis

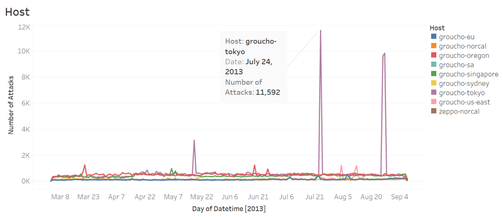

Fig. 6 Identifying the Host

Timeline analysis on determining the Host.

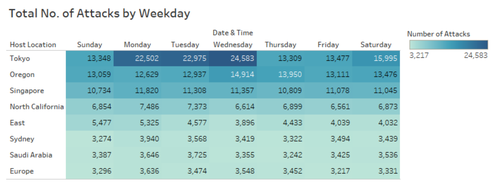

Fig. 7 Number of attacks by week

Timeline Analysis is used to discover deeper information/trends about the dataset regarding the number of attacks over time. Time is further broken-down into weekdays and the initial results presented on the heatmap indicate the trend that attackers seems to be more active on Monday – Wednesday as compared to Thursday – Sunday.

Deeper analysis on Tokyo

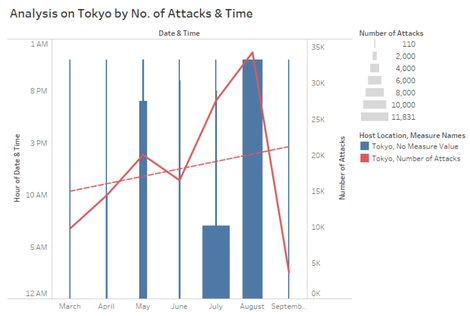

Fig. 8 Timeline analysis on Tokyo

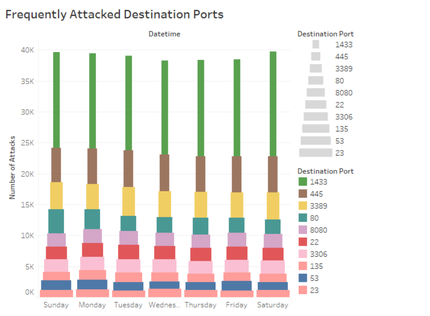

Fig. 9 most frequently attacked destination ports

Figure 9 indicates most frequently attacked destination ports and the number of attacks during the week.

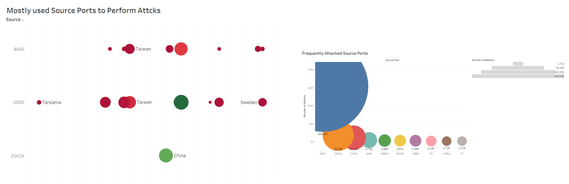

Fig. 10 Mostly used source ports to perform attacks

Conclusion

This project results imply the efficiency of Honeypot and how it can be used as a security mechanism to fight against cyber threats. Major benefits of honeypot can be listed as;

- Identify new and emerging malware[5]
- Functions as an early warning system[4]
- Identify source attack
- Spot insider Threats
- Confuses Attackers
- Streamlines security alerts