Tu Lam

CS 373 (Defense Against the Dark Arts)

Dr. Bram Lewis

November 15th, 2021

Homework #8 Write-Up

Week #8 of the class have a special guest; their name is Eric Peterson. They work as Research Manager at McAfee. Through this week lecture, we will be dealing with the ***messaging security*** and it will be divided into two parts. The first one deal with the history of the ***anti-spam landscape*** and related terminology. While the second part deal with the aspects of ***email-borne threats and basic data manipulation*** techniques for identification and automation of human classification/content creation.

First, we dive into the ***terminology of messaging security*** and below are the term and their definition of it:

1. **Spam / Ham**: *Email or messages that are bad to user (Spam), while there are good one that does not attack user (Ham).*
2. **Spamtrap / Honeybot**: *This is a tool to collect spam*.
3. **Botnet**: *A number of Internet-connected devices, each of which is running one or more bots. Botnets can be used to perform distributed denial-of-service attack, steal data, send spam, and allows the attacker to access the device and its connection*.
4. **Snowshoe Spam**: *To technique spam use to go over different IP address so they get minimize chances of being filter as "spam" when receive by user*.
5. **Phishing vs Spear Phishing**: *Fraudulent attempt to obtain sensitive information such as usernames, passwords and credit card (Phishing). While the other involves interacting with a targeted individual to gain personal or other confidential information (Spear Phishing)*.
6. **RBL**: *Also known as Realtime Blackhole List, a list of IP addresses whose owners refuse to stop the proliferation of spam*.
7. **Heuristics**: *Various algorithms and resources to examine text or content in specific ways. In this case a way to rank and rate the level of spam*.
8. **Bayesian (Statistical)**: *A tool to recognize certain words and the likelihood that they're related to spam*.
9. **Fingerprinting / Hashing**: *Combining of two methods to easily track down the origin of spam*.

Then, we move onto the ***history of spam*** where we first dive into the classic of ***419 phishing***. This number 419 comes from the international phone code that makes this illegal, also called the ***Nigerian Prince Scam***. And this is the start of most of phishing email come from to the origin of it. The most common things you receive from spam is asking you about your personal data, transferring money, or act like a close family member or friend to access any data from you that they are not supposed to access. Another famous classic phishing is ***Canadian Pharmacy***, where this spam is sending user a promotion in pharmacy throughout Canada to buy low price medication via this website, but they actually take the money instead of selling medication. There more examples of the phishing things out there, but there was also a rise in ***Botnet*** attack out there in the cyber world. Below is an image of the history of Botnet throughout the year.

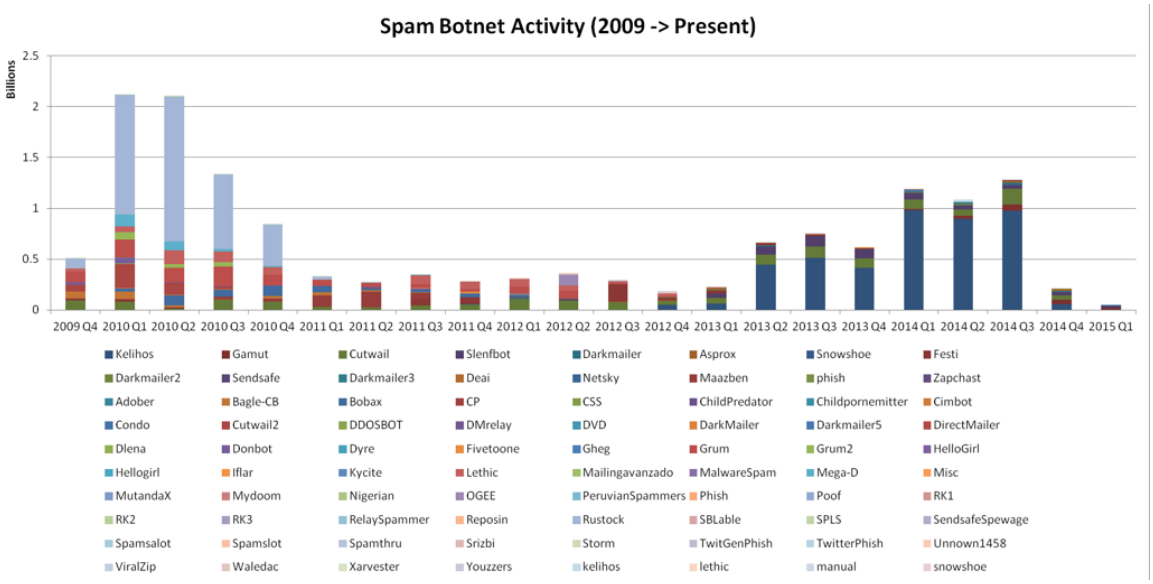

Figure #1: [**From Lecture**] *The layout of Botnet Spam History*

To the next topic, we discuss ways to combat spam by providing two methods, ***reputation-driven*** and ***content-driven***. With reputation-driven, you have the ***IP***, ***message***, and ***URL*** while content-driven you got ***string***, ***regular expression***, ***message attributes***, and ***meta rules***. There are also ***tools*** beside the method to help us out identify the spam such as: ***Linux tools, open-source database, regex coach, trustedsource.org, spamhaus.org, and more***. Then we have ***research techniques*** where they give us tool to deal with data flooding in when it comes to phishing and identifying them.

Finally, moving onto the part 2 of lecture, we are wrapping it up and look at ***SMTP conversation***. This is a way to identity how email is communicated through network and look to see if this is a spam or ham. Lastly, we look at the ***data scientific method*** where we start out as ***getting the data, then develop intuition about the data, then formulate the question, leverage the data, create a framework, and analyze the result***.

Overall, the topic for this week deal more about the history of phishing and the content of it. Then we look at tools that can help us identify these phishing things that people tend to take it as it is an official email from somebody. With these tools and methods in mind, we can learn way to identify phishing items and protect us online from harmful things.

References

1. Peterson, E. (2015, February 24). Messaging Security. Canvas. Retrieved November 15, 2021, from https://canvas.oregonstate.edu/courses/1877198/pages/module-8-learning-materials?module_item_id=21556825.