Determining Selection Behavior of Malicious Bots within Twitch Chat

Andrew Yarbrough, Tanner Bonds
asy0003@auburn.edu, tjb0057@auburn.edu
Auburn University

Abstract

Spam attacks, phishing attacks, and many other forms of attacks that cause users to go to malicious links so that attackers can steal data from people are very prevalent in our technology-centered society. We see attackers posing as banks, or other 'trustworthy' organizations send emails to individuals asking for their bank information and people giving this information up because they are none the wiser. Many corporate organizations will spam social media accounts, trying to get individuals to click on their links, so they get paid for every click that tracks to their website. One platform that also deals with these same issues that are continuously growing is the game-streaming platform. In this paper, we talk about one such platform called Twitch. Twitch does little on the server-side to handle the posting of malicious links, which leaves all the work to the streamers. Our goal for this paper is to determine the selection behavior of malicious bots and 'spam' links to lead to people promoting themselves or some other entity. We investigate each link for malicious content using Google's safe browsing API[1.]; then, we look at the remaining links and determine how often they occur (Spam attacks) and what their contents are to see if they're promotional links (Form of phishing attacks). We expect there to be many links that contain malicious content, but, as you will see in our results, that was not the case.

1 Introduction

Twitch is a popular streaming platform that hundreds of thousands of users use daily. As the growth of technology continues, we see more and more people watching Twitch. According to TwitchTracker, as of the writing of this paper, Twitch has already gathered 243 Billion minutes of watch time from viewers [2]. As Twitch continues to grow in numbers, more and more people will want/continue to exploit these numbers by posting any malicious links they can. These malicious links could contain viruses, malware, propaganda, etc. that many people are not aware of and may have a difficult time noticing when someone is posting a malicious link.

The primary purpose of our research is to determine the selection behavior of malicious bots within Twitch. Before we outline the approach we plan to take, we will first give some background information about what specific elements we are looking for within the chat logs, what Twitch is and its importance on the internet, and further define our topic and better define what we mean by malicious.

The primary characteristics of these malicious bots are spam and phishing attacks. Spamming is the act of sending unsolicited messages, often repeatedly, on the same website. These spam messages are typically used for advertising purposes. Within the context of the Twitch platform, bots, and occasionally users, will post links within the chat log of a streamer. These links are usually directed to some product or advertisement. While an average user can implement these spam attacks, they are more commonly sent by bots, as this allows for automation and an increase in the number of messages that can be sent in a certain amount of time. Phishing is considered to be an attempt by a fraudulent entity to obtain sensitive information such as passwords, credit cards, or usernames, specifically if the fraudulent entity is masquerading as a trustworthy individual or body. As with spam, bots and users can implement these attacks within Twitch chats by sending messages or links appearing to be authentic in hopes that a user can be coerced into entering valuable information.

Twitch is a video live-stream service that focuses primarily on users playing video games, but can also include other broadcasts such as eSports competitions, music, and other creative content. By February 2014, Twitch was considered the fourth-largest source of peak internet traffic in the United States. As of May 2018, Twitch had an average of 2.2 million broadcasters and 15 million daily users. The content on Twitch can be viewed as video on demand, but most content is streamed live. There are primarily two types of users on Twitch, viewers, and streamers. Streamers are those that broadcast their content, such as themselves playing video games, and the viewers are all others that watch their content. Within each streamer's channel is a chat system, in which those viewing a stream can comment in real-time, with both the streamer and all other viewers being able to see such messages. It is this system that malicious attackers choose to exploit.

A bot is any sort of user that runs an automated task, also known as a script, over the internet. Bots typically perform tasks that are simple and repetitive but can complete these tasks at speeds of a much higher rate than humans can, all while expending less resources. More than half of all traffic on the web is composed of bots[3]. Bots can be used maliciously by coordinating and operating automated attacks on networked computers. These bots can be used for many nefarious applications, such as DDoS attacks, programs that harvest bandwidth by downloading entire websites, and many more. Specifically, the type of bot that we are focusing on is spambots, bots that post links that attempt to redirect users to malicious websites.

Throughout the rest of this paper, we will provide a planned approach to research the behavioral selection pattern given to these malicious bots. First, we will provide related works from Twitter about spam and phishing attacks and how it relates to our research in this paper. Then, we will discuss our approach to gathering data and analyzing the data gathered from Twitch to determine the selection behavior behind the bots. Finally, we will close the paper by providing a short section that leads to this type of research being expanded on in the future and a conclusion that summarizes the preceding sections.

2 Related Work

The topics of phishing and spam are not new on the internet. Many studies in the past have looked at phishing and spam attacks on many different platforms, such as social media, email, text messages, and many more. In this paper, we will be looking at comparisons between our findings on Twitch to past findings on Twitter. Some related works on Twitter have looked at phishing attacks on the platform in great detail. One such paper by John Seymour and Philip Tully, talk about a program/application called [SNAP-R] that targets groups of users on Twitter and sends them appropriate phishing tweets based on machine learning techniques [5]. The tool takes general phishing tweets as the learning component of the neural network and then models its own tweets off the learning component. Then, the machine learner will create groups and categorize individuals into these groups based on their profile, what organization they adhere to, and what time of content they tweet or retweet on Twitter. From there, the machine learner then creates different bots for each category to prevent being detected by Twitter's ToS (Terms of Service) and send relevantt' tweets to these groups in order to get people to click on the links in the tweets. The paper revealed that [SNAP-R] had a 30-66% success rate in its phishing attacks. Another paper by Mohammad Shafahi, Leon Kempers, and Hamideh Afsarmanesh do a similar study with bots that post tweets about a specific subject for four weeks based on trends from people who click on phishing links [6]. The paper revealed that 437 unique users could have been phished based on the trends and 33 visited the links provided by the bots.

Moving on from phishing attacks, many papers in the past of looked at spam attacks on Twitter as well. The purpose of spam is to reach as many individuals as possible and to bombard users with information to get them to click on, or invesitgate, the information. A paper by Cristian Lumezanu and Nick Feamster shows how companies can use spam on yahoo's email and Twitter's social media platform to spam a larger quantity of users and get more people to click on their sites [7]. Another paper by Grant Stafford and Louis Lei Yu looks at how spammers make use of trending topics in Twitter to produce large quantities of spam tweets to acquire more users to click on their links [8]. In both papers, researchers look at how spam is used and what spammers look for in order to gain a large quantity of users that will produce more hits on their links.

From both of these topics, we can draw many comparisons from past works to our topic with Twitch. A vast majority of links in our data are links that are posted to phish users into clicking on them, and these links are posted several times throughout a stream by the streamer. Both the purpose of the links and the repetition of them being posted that phishing attacks and spam attacks are very prevalent on the platform. Section 4.4 further talks about our comparison between Twitch and Twitter's spam and phishing attacks.

3 Approach

In this section, we will outline our planned approach to research the selection behavior of malicious bots in Twitch thoroughly. Our approach will consist of three major phases: data gathering, analysis, and comparison. We will list out the designated time from each section and go into detail about what will occur in each phase.

3.1 Data Gathering

We spent a week and a half gathering data. The time frame that we chose to spend gathering data is from March 18th to March 29th. These weeks were chosen due to their relative overlap with several Spring Break weeks. We believe this time frame will have a higher volume of malicious bots being present in Twitch chats due to a high volume of users that will be using Twitch during their spring breaks. Table 1 below shows how much time viewers spend watching Twitch grows during March compared to other months that occur during school semesters [2].

Table 1: Twitch Statistics

Month	Avg. Concur. Vwrs.	Avg. Concur. Chan.	Time Watched	Active Str.
May 2019	1,262,561	48,294	939M hrs.	3,933,404
April 2019	1,234,112	51,261	889M hrs.	4,184,167
March 2019	1,273,772	59,934	947M hrs.	4,389,867
February 2019	1,311,802	56,286	880M hrs.	4,205,001
January 2019	1,276,009	55,794	949M hrs.	4,537,807

Stats pulled from a Twitch statistic's tracker website https://twitchtracker.com/statistics (Concur. = Concurrent, Vwrs. = Viewers, Chan. = Channels, Str. = Streamers)

In order to gather the chat log data we need, we will be using a Twitch Chat Logger program found on GitHub [3]. The Chat Logger does not have any functions that write to the Twitch chat of any channel it enters. To ensure it does not do this, we will thoroughly review the source code and make any changes to ensure that the bots do not type in chat. Also, we will thoroughly test the chat logger in the weeks prior to data gathering to ensure that the Chat Logger does not post any errors in the Twitch chat or from any other source of errors.

3.2 Methodology

Each of us ran ten instances of the Chat Logger on our machines, for a total of twenty channels monitored per run. Ten Chat Loggers were monitoring the streams with the Top Ten highest view count on Twitch at the time that the runs began. The other Ten Chat Loggers were monitoring the Top Ten highest view count below 1,000 viewers. The Chat Loggers ran for several hours at a time before the runs were terminated, and the channels being monitored were updated to the current Top channels. The Chat Logger stores the data gathered from Twitch will be stored in *.log* files. A single *.log* file is generated per run of the Chat Logger.

3.3 Analysis

Once we have gathered all the data during the one-and-a-half week period, we will spend the following two to three weeks analyzing the data. The main goal of our research is to determine the selection behavior of malicious bots in Twitch; to do so, we must first find any malicious entities. With all of our Chat Logs gathered, we will filter them through a program to remove all messages that do not contain links within them. Once this is complete, we will enter these links into Google's Safe Browsing Check API. The purpose of this is because users may post links that are images or clips that the streamer allows to be posted in the chat. We want to remove as many of these non-malicious links to narrow down our search further. Once we have our links narrowed down to the type of malicious bots we are looking for, we can begin to analyze the malicious bots through all the logs to determine any patterns. Patterns we are looking for are similar bots with similar links across multiple channels, different bots with similar links across multiple channels, and similar bots with different links across multiple channels. Looking at these three different types of patterns, we will be able to see which types of streamers and what parts of Twitch these malicious bots go to in order to get as many people to click on their links as much as possible.

3.4 Malicious Bots on Twitch and Twitter

Once we have gotten all of our data, we plan to compare our findings of malicious bots on Twitch with similar phishing and spam attacks on Twitter. One component we want to look at is how the links are formatted. Twitter limits tweets to 140 characters; thus, individuals who want to post links need to use URL shorteners. Spammers have also adopted the URL shorteners to camouflage and improve the user click-through of their spam URLs [4]. Phishing attacks on Twitter are also a concern. Phishers have started using Twitter as a medium to spread phishing because of this vast information dissemination [5]. Once we figure out the selection pattern of malicious bots on Twitch, it would be interesting to see how it compares with malicious bots, or accounts, on Twitter as well.

4 Results

From Twitch, we managed to log a total of 11,508,772 messages from 286 different channels, shown in Table 3. When we filtered these messages to find those that contained only links, we found that only 119,207 messages had contained an HTTP or HTTPS link, meaning that only 0.01% of all messages pulled contained links. Furthermore, many of these messages contained duplicate links, i.e., youtube.com, google.com, and others. We then filtered these links to remove duplicates and returned a total of 12,483 links. The following subsection will discuss our findings of malicious links we found in these 12,483 unique links.

Table 2: Chat Totals

Top 10 Chat Msg.	Top 1K Chat Msg.	Total Msgs.	Total Strs.
10,738,038	770,734	11,508,772	286

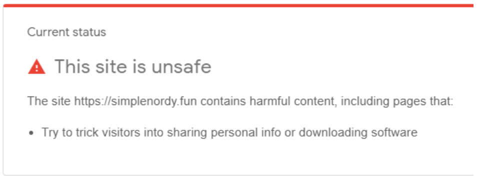
4.1 Malicious Links

Upon entering our links into Google's Safe Browsing API, we found only four unique links that were reported as malicious. These four links occurred 203 times in one stream by a "nightbot," a bot that is usually created by the streamer, and once by a viewer, shown in Table 4.

Table 3: Malicious Links

Links	Occurences	Date	Timeframe
https://simplenordy.fun	103	3-24	11:14:55 AM - 11:46:27 AM
https://simplehot.fun	99	3-25	08:06:50 AM - 09:00:34 AM
https://simplehot.fun?referral=5359	1	3-25	8:10:09 AM
https://simplehot.fun?referral=8645	1	3-25	8:17:23 AM

Upon investigating why Google returned these four links as malicious, we discovered that Google deemed these links as malicious because the site contains harmful content that may be trying to trick visitors into sharing potential info or downloading software. Figure 1 below shows a screen capture of the message google presented us for one of the four links.



Current status

This site is unsafe

The site https://simplenordy.fun contains harmful content, including pages that:

- Try to trick visitors into sharing personal info or downloading software

Figure 1: Google Report for Malicious Link.

We believe that these links are not malicious. Google's Safe Browsing API is designed to look for specific characteristics that are commonly found in malicious sites, like the ones shown above; however, we think the links that the streamer posted were simple clips or a small mini-game that the streamer was posting for his/her viewers to be interactive in chat. Without running the risk of corrupting our machines, we can not investigate the link thoroughly as we do not have the equipment or environment to open the link personally; so, we will count the link as malicious according to Google's API.

4.2 Advertising Links

We stated that we found 12,483 unique links, shown in Table 4, and reported on 204 of them. So, what about the other 12,279? Google's Safe Browsing API did not return these links as malicious; instead, they were found to be safe to travel too. However, as we stated earlier, we are also looking for spam and phishing links. We investigated the remaining links to see if there were any trends in them and what type of links they were and found that each link was either a self-promotion link, a link promoting another streamer or clips that streamers allow viewers to capture of the stream for replays. Streamers make money from their streams by having a large amount of follows under their wings. So, in order to keep their numbers high, streamers will post promotion links that either lead viewers to join their messaging platform (Discord), or their social media accounts, or even sponsor links. Sponsor links are links provided by sponsors to streams to promote their product.

Table 4: Total Links

Top 10 Chat Links	Top 1K Chat Links	Total	Total Unique Links
102,241	16,966	119,207	12,483

An example of a sponsored link would be Rogue Energy Drink providing a link to a stream for them to post in their chat to promote the drink. In exchange, the sponsor pays the streamer a large sum of money (usually in the thousands) passed on their viewership. We also mentioned that some of the links are for short highlights videos, commonly called replay clips. Many streamers have a unique bot/program in place on their streams that allow viewers to type a particular command in chat that will record a certain amount of time in the stream and then replay it. The times vary depending on the streamer's settings.

4.3 Streamer Security

As we have shown, less than 1% of the messages we logged contained any form or link. While this is not what we were expecting, it is excellent news! Twitch does not handle malicious security on the server-side. Instead, Twitch has a reporting system that allows viewers and streamers to report behavior that does not adhere to Twitch's rules and policies. However, this system is flawed. Once someone is reported, they are not immediately banned or blocked from using Twitch. It takes days for a report to be evaluated by Twitch's moderator team before determining what to do with a reported user. Thankfully, Twitch gives the control of how to protect a streamer's channel to the streamer. Many streamers use some of the default settings that Twitch has in its UI to prevent viewers from posting any links. For example, we asked a few streamers how they handle viewers posting links in their chat, and they said that Twitch has a setting that allows them to prevent anyone other than themselves from posting links. We also watch Twitch streams in our spare time and noticed that some viewers can still post links. We asked them how they handle viewers posting malicious links in their chat, and they said they have a handful of personal moderators who monitor their chat and delay any links from being posted. Their mods will investigate the links and make sure they are safe for posting before it is allowed to be posted. If the link is malicious, they will ”shadowban” or ban viewers from the chat. Shadowbanning is a timed ban that allows the viewer to keep watching the stream and type in chat; however, no one else can see what they are typing. We were not expecting the streamers to be aware of the possibility of attackers trying to post malicious links in their chat or even bother with preventing them. Thankfully, streamers are a lot more conscious of their viewers and look out for their best interests as well as their own.

4.4 Twitch and Twitter

In regards to spam and phishing attacks on Twitter as opposed to Twitch, are results are inconclusive. However, this in and of itself, presents key findings. We find that there is little to no malicious presence on Twitch. Twitter, however, has been plagued by bots conducting spam and phishing attacks, spreading misinformation, and other malicious actions practically since its inception. Twitter was released in March of 2006, whereas Twitch was released in June of 2011, almost a five year difference. While Twitch mas maintained a relatively bot-free in terms of maliciousness, Twitter continues to be a platform where bots can thrive, with little to no intervention from site administrators.

5 Future Work

All of our work was centered around the streaming platform Twitch and comparing phishing and spamming tactics with Twitter. However, there are many other streaming platforms than Twitch. Such streaming platforms are Mixer (Microsoft), Caffeine.tv (Third-party), YouTube (Google), and many more. Our work could be expanded on to see if there are any differences between the platforms when it comes to handling security against malicious links. Another path to expand on for this project is to see how the bots/mods that are created/appointed by streamers monitor chat for links being posted and to see how secure these scripts are. We find vulnerabilities in security protocols, tools, and policies every day; so, it would be interesting to see how secure these bots/mods are and develop better ways to improve them.

6 Conclusion

In this paper we have described a problem that has plagued social media sites for years. We hoped to analyze chat messages from the popular Twitch website in order to find malicious links, specifically posted by bots, to determine how these bots choose their targets. While we hoped to find a significant amount of links, what we instead found was a much better case. We found that their were an insignificant number of links being posted, and that the users of Twitch appear to have taken matters of security in regards to malicious links seriously. While disappointed that we were unable to find an answer to how these targets are chosen, we are much happier with the outcome.

References

[1] ”Google Safe Browsing”, Safebrowsing.google.com, 2020. [Online]. Available: https://safebrowsing.google.com/. [Accessed: 25- Mar- 2020].

[2] “Twitch Statistics & Charts,” TwitchTracker. [Online]. Available: https://twitchtracker.com/statistics. [Accessed: 18-Feb-2020].

[3] M. Santos, ”melsantos/Twitch-Logger”, GitHub, 2019. [Online]. Available: https://github.com/melsantos/Twitch-Logger/. [Accessed: 25-Mar-2020].

[4] Click traffic analysis of short URL spam on Twitter - IEEE Conference Publication. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6679991. [Accessed: 19-Feb-2020].

[5] J. Seymour and P. Tully, Tutorial.evogtechteam.com, 2020. [Online]. Available: https://tutorial.evogtechteam.com/wp-content/uploads/2017/03/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter-wp.pdf. [Accessed: 17- Apr- 2020].

[6] M. Shafahi, L. Kempers and H. Afsarmanesh, ”Phishing through social bots on Twitter - IEEE Conference Publication”, Ieeexplore.ieee.org, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7841038. [Accessed: 17- Apr- 2020].

[7] C. Lumezanu and N. Feamster, ”Observing common spam in Twitter and email — Proceedings of the 2012 Internet Measurement Conference”, Dl.acm.org, 2020. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/2398776.2398824. [Accessed: 17- Apr- 2020].

[8] G. Stafford and L. Lei Yu, ”An Evaluation of the Effect of Spam on Twitter Trending Topics - IEEE Conference Publication”, Ieeexplore.ieee.org, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6693355. [Accessed: 17- Apr- 2020].