Beware the evil bots: e-commerce thieves and spreaders of "fake news"

Linda A. Bressler
Southeastern Oklahoma State University

Martin S. Bressler
Southeastern Oklahoma State University

ABSTRACT

Technology is an integral part of everyday life, whether at work, school or home. Technology makes our lives' easier in many ways. However, technology can also be used for illegal and unethical activities such as for spying, theft, and even spreading false information across the Internet. In this paper, the authors describe how bots are becoming one of the leading-edge tools of e-commerce and online information exchange. The authors discuss various types of bots and identify available security counter-measures to defend against attacks.

INTRODUCTION

The concert you have been waiting for is finally announced to be at your favorite local venue. As soon as ticket sales open online, you are at your computer searching for the best seats you can find. However, within seconds, all the best seats are sold out. You desperately want to see this musical group, so you order tickets through a ticket broker, but instead of the stated venue price of forty dollars, you now must pay ninety dollars. How could all the best tickets be bought up by brokers within seconds? Blame it on the bots!

Bots, (short for robots) or in this instance, “rush bots,” jump to the head of the ticket line to buy up all the best tickets for resale. This may be considered enterprising for the bot-users, but frustrating for consumers. Some bots assume more malicious tasks such as sabotaging a competitor’s online advertising or spreading “fake news” on the Internet.

The problem is that this technology, like so many other technological developments, can have both legitimate commercial use and at the same time, the potential for both illegal and unethical applications. Bots can be used to sabotage a competitor's advertising efforts, to overcharge customers for online advertising, to drive up the cost of brokered products such as concert tickets, and to social engineer voter's opinions on political issues or candidates.

Bots are one of the latest technological developments to impact our online lives. E-commerce is not the only online presence for this new technology. Bots have moved more into the spotlight in the last few years as they attempt to control online news and discussions. So, what exactly are bots and how can we best manage this technology to productive uses?

According to *Business Insider* (cited in Bonderud, 2017), 80% of organizations already use these automated response devices or intend to put them into operation. Likewise, individuals have integrated bots with their social media technology for a variety of purposes. As an example, Douglas (2018) points to Rob Manuel’s *Smash Hits Interview Bot*, which mimics the open-style interview questions of the 1980’s magazine *Smash Hits*. With the help of this bot, social media users can interview themselves and create a funny personal interview.

DISCUSSION

Bots defined

Bots are automated software programs that can execute specific commands when it receives a certain input (like a ro-"bot"). Bots are most often seen at work in the Internet-related areas of online chat and Web searching. The online chatbots do things like greet people when they enter a chat room, advertise Web sites, and kick people out of chat rooms when they violate the chat room rules. Web searching bots, also known as spiders and crawlers, search the Web and retrieve millions of HTML documents, then record the information and links found on the pages. From there, they generate electronic catalogs of the sites that have been "spidered." These catalogs make up the index of sites that are used for search engine results (https://pc.net/glossary/definition/bot).

Put more simply; bots can be considered an application that utilizes artificial intelligence to initiate human conversation. The potential for improved customer service could be significant. For example, Bonderud (2017) points to the National Health Service in the United Kingdom, which developed a chatbot app which will ask patients illness symptoms and then based on patient response, suggests professional medical assistance or alternative solutions. The patient

chatbot could eventually replace the National Health Service nonemergency call line, resulting in saving millions of dollars.

Gilchrist (2017) reports that chatbots currently save businesses about $20 million per year globally. That number is expected to $8 billion per year by 2022. Maxim Abramchuk, Founder and Chief Technology Officer at BotCube (cited in Gilchrist, 2017), says that an automation chatbot for a large company can cost anywhere from $20 thousand to more than a million dollars.

Spammers use Bots to infiltrate and attack online content for a variety of commercial and social objectives. Spammers can attack across a wide range of targets including blogs, social bookmarks, Wiki, video sharing, online forums, online communities, opinion/review sites, email, instant messaging, VoIP, mobile phones, and search engines. In effect, spammers use any communication used by humans and deploy whatever means necessary to attract potential victims.

Some experts (Liu et al., 2014) consider bots and botnets to be one of the most severe threats to Internet security in recent years. When compared to other malware methods such as worms and viruses, bot behavior can be especially stealthy. This stealthy-ness makes bot detection extremely difficult. For instance, a bot can remain inactive without significant activities for an extended period. In most situations, a bot only generates a small amount of traffic, which then remains hidden surrounded by legitimate traffic.

According to Liu et al, typically **"**a bot will exhibit three invariant features along its onset: (1) the startup of a bot is automatic without requiring any user actions; (2) a bot must establish a command and control channel with its bot-master; and (3) a bot will perform local or remote attacks sooner or later. These invariants indicate three indispensable phases (startup, preparation, and attack) for a bot attack" (2014).

How much do bots cost? Bonderud (2017) cites a report in Venture Beat that indicates whether you develop the bot in-house or contract with an outside firm, the initial setup costs would typically range from $5,000 to $10,000 in addition to monthly maintenance fees of around $2,000. This allows even small enterprises to deploy multiple bots to accomplish various tasks.

Bots used in e-commerce

Millions of individuals and companies utilize the Internet via e-commerce including small business. Some of the more popular gateways include Amazon.com, eBay.com, as well as most retail outlets such as Macy's and CVS. As this can be the most profitable way to sell products, more sellers will be attempting to attract the attention of prospective consumers. One way to keep existing customers and persuade new customers to buy a company's product would be to offer suggestions to potential buyers by using an "intelligent bot." This could be even more productive for the e-commerce companies by tracking or retrieving information from the customers by the intelligent bot (Shahmanzari & Okzan, 2014).

Rush bots used in ticket sales

Bhave & Budish (2017) noted that in early 1868, a reading by Charles Dickens in New York City tickets sold for $2.00 and a secondary market offered tickets for $20 and even more interesting could be a report that a child paid $30 in gold to sit closer to Mr. Dickens during his

reading (*New York Times,* December 1867). As can be noted with the Dickens reading, secondary ticket sales markets have been operating for many years.

In a recent report (Independent Review, 2016), concerns regarding manipulation and ticket fraud were raised in Parliament when the Consumer Rights Bill (CRA) passed, but the CRA aimed only toward the secondary ticket process and better transparency. Schneiderman (2016) stated that the New York Attorney General continues to receive complaints of being unable to purchase tickets even minutes after the tickets are released to the public. It appeared that this was not the case of supply and demand issues, but prospective buyers believed that the ticket-buying process was fixed. The report also reminds the reader that ticket brokers sell accumulated tickets at sometimes 1,000 percent of the original cost of the tickets and that fees can be assessed at over twenty percent of the ticket price.

The way the ticket process works would be upon release of the tickets, the brokers buy up the maximum amount of tickets possible and brokers will often use bots to purchase tickets at high speeds leaving only a few for sale (left over after reserved tickets for artists, promoters or various industry elites or in-group individuals). The report calls for greater transparency, perhaps requiring promoters of events to make transparent to the public the number of tickets available to the public (Independent Review, 2016). The report also suggests that ticket sellers who use bots to purchase vast amounts of tickets should face criminal prosecution. Finally, the report indicates that New York should initiate a reasonable dollar limit on resale ticket markups.

Ad-click bots

In 2014, a Mercedes-Benz online advertising operation noted that just under 50% of the ads resulted in human viewing. Google later announced that 56.1 percent of ads on the Internet would not be displayed. Neal & Kowenhoven (2015) reported from their research that ad fraud could be considerably higher than was initially thought. Neal & Kowwenhoven (2015) conducted a study aimed to compare the ratio of Ad-Clicks initiated by humans to computer-generated ad-bot clicks, which resulted in some fascinating findings. The authors noted the prevalence of spambots with their purpose to spam vast amounts of content on the Internet and programmed to add advertising links at the same time.

Also, bot farms will often be utilized with online app stores such as Google Play for manipulation purposes to increase positive ratings and reviews. However, even worse will be the use of bots for an attack on network computers for denial-of-service attacks. The authors cite a 2012 report by CNN reporter, Percy Lipinski, who provides the description below of Bot behavior:

Commonly Observed Bot Behavior

All bots have a common set of properties
Bots primarily exist, directly or indirectly, for economic gain
Mimics, to any extent, the actions of a human using a computer
Repeats such actions multiple times
Initiates activity

Bots, combined with artificial intelligence, have tremendous commercial potential for reducing business costs and increasing profits, in addition enhancing the personal use of technology, most notably social media.

Bots are designed only to carry out the necessary minimum actions required to complete its task. Bot behavior, at the atomic level, falls into any one of the following classifications (with examples of type):

1. Sends a single message (Denial of Service Bots, Distributed Denial of Service Bots, Ad Click Bots, Ad Impression Bots)
2. Sends a single message and waits for a response (Email Spam Bots, Ad Click Bots, Ad Impression Bots, Online Banking Bots),
3. Sends multiple messages asynchronously (Denial of Service Bots, Distributed Denial of Service Bots),
4. Sends multiple messages asynchronously and waits for one or more responses (Online Spam Bots).

Note that in behaviors 2 and 4, the sender address (i.e., the IP address) must be valid for the response to be received (although not necessarily the point of origin), while behaviors 1 and 3 can accomplish their task without this prerequisite condition making them considerably harder to detect their actual position of origin.

These bots can badly damage companies' e-commerce sales. Recently, a Nevada Hyundai dealership won their fight with a bot's advertising attack. The marketing manager at this dealership extensively utilized digital advertising and found that many of the clicks on their advertising link ended up being bots…not humans, that did not help the company's advertising endeavors. The manager used a bot-detection service called Orbee which investigated the source and intent of the company's advertising traffic to investigate that his advertising dollars would be reaching humans, rather than bots. The authors noted that because of the marketing manager's investigation, he could prove the traffic he paid for was tainted or out of the market and he was able to negotiate contracts with vendors not fulfilling the company's needs. In one instance, after a conversation with one vendor explaining his advertising dollars were being wasted on bots, one vendor "re-targeted" its' advertising for the dealership because 55 percent of the traffic was unusable bot traffic.

Unethical social bots

Salge and Berente (2017) defined Social Bots as "computer algorithms in online social networks." The authors indicated they could be used to share messages and pictures while connecting others to the Internet. Ferrara et al. (2016) noted that bots are not new and could be identified since the beginning of the Internet. Bots, initially utilized for good reasons, now can be involved in social engineering endeavors, political goals, ticket resales (Cui, Duenyas, & Sahin, 2014) or as attack web bots.

Bots as social engineers

Web bots can be used to steal personal information, spread "false news," manipulate and deceive social media users. Web bots can be used to gather information from many sources (Facebook entries, news outlets, etc.). If the initiator wished to deceive social media users, he/she would design algorithms that emulate human communications such as utilizing a web bot by attempting to spread political misinformation intending to sway voters before a coming election. This happens during an attack when the bot re-tweets posts without checking that the post is accurate, nor testing the credibility of the tweet source (Attkisson, 2017, 126-129).

Another problem can be when the bots are programmed for malicious means whereby the (for example) tweets deceive, misrepresent, fabricate or even, exploit with the use of spam, rumors, malware, smear, slander, defame or perhaps also flooding a user or users' account with hundreds of unwanted and sometimes frightening and threatening messages (Ferrara, et al., 2016). Attkisson (2017) reports that these attacks can be as simple as disagreeing with someone who posted a political opinion with hundreds of re-tweets to influence political views or smear political candidates.

Bot counter-measures

Hayati et al. (2011) identified several of the most common anti-robot systems, or countermeasures, which are listed and described below in Table 1. As noted, some of these techniques do not entirely stop spamming, but instead, only slow down and make spamming even more costly to spammers.

Perhaps one of the most effective and widely-used means to combat bots is a Completely Automated Public Turing test to tell Computers and Human Apart, or CAPTCHA. CAPTCHA is a challenge-response technique requiring users to type in letters and numbers into a form. On the downside, as computer programs get better at deciphering the CAPTCHA, the effectiveness of CAPTCHA will quickly decline.

One of the current approaches to detect web Bots is through monitoring web usage data. Monitoring web data includes tracking the IP address access, web page access, and browser use (Mozilla Firefox, Internet Explorer, etc.). By monitoring behavior, clear patterns develop over time, making it easier to profile spambots. Web usage trackers can also track where and how spammers navigate through the web.

Table 1 Bot counter-measures

Counter-measure	Description
IP Address	well-known robot IP addresses are listed on certain websites and can be easily retrieved to help detect web spiders
TXT Files	a text file of access restricted places within the website that forbids web robots from accessing the entire website
User-agent	identification method to determine and approve the application for access to the server
Head request	head requests check the validity of hyperlinks, accessibility, and recent modifications to the requested webpage
Referrer	the Referrer identifies whether the requested is generated by a human or by a spambot
Flood Control	flood control limits the number of requests a client can send within a specified time frame. As spambots attempt to flood a website, the flood control countermeasure prevents the spammer from gaining website or file access
Nonce	uses a random set of characters placed on a webpage with a form used to prevent automated form submission
Form variation	form variation is like Nonce, except that form variation uses a varying set of objects (such as click on all the pictures with street signs) rather than characters
Hashcash	like Flood Control, Hashcash requires the sender to calculate a stamp for access. Calculating the stamp becomes costly for the spammer, and although it does not stop spamming, it often slows down the amount of spam

Spammers usually obtain email addresses by stealing them from various websites. The spammers use "spambots," or computer programs to routinely troll web pages to steal email addresses. Project Honey Pot helps to identify spambots by providing unique tracking addresses for your website (https://www.projecthoneypot.org/how_to_avoid_spambots.php).

"Munging" is another technique used to defend against spambots using one of two strategies for address munging. The first strategy is to redesign the addresses on your website so that they are invalid for use by the spambot but can be easily fixed by humans. The second method is to conceal addresses on your web pages so that the spambots will be unable to locate them. Each of these techniques is discussed below.

An example of the first method would be to make your web address technically invalid by inserting random text that spambots won't be able to recognize as not being part of the

address. However, humans will understand that they need to remove the text before sending to you. In the example below, the same address is expressed in three different ways:
carl@**REMOVETHIS**example.com
carl**DELETEBEFORESENDING**@example.com
Zarl@example.**Z**om (replace **Z**s with **C**s)

Spambots will normally gather these concealed web addresses, but when spammers attempt to send messages to them, the messages will not go through. Unfortunately, this creates additional traffic on the network as well as on your mail server. Perhaps more troubling, legitimate visitors to your website will incorrectly "de-mung" your address and will therefore not be able to send messages to you.

The second strategy is based upon concealing web addresses from spambots to stop them from being collected. Should you choose to hide addresses from spambots, it is important that you understand how spambots work. Spambots typically find addresses by searching for configurations of text that resemble email addresses. Email addresses, for example, contain an @ symbol. Therefore, spambots scan webpage text in search of any @s. As indicated below, if you eliminate the @ from the web addresses, then most spambots won't be able to recognize that your addresses:
carl**-at-**example.com
carl**(at)**example.com
carl **AT** example **DOT** com

Although this method effectively hides your web address from spambots, visitors to your site will usually still incorrectly de-mung your address, or not even recognize it as an email address, and therefore, be unable to contact you.

A more complex adaptation of concealing your address still allows human users to view the addresses without obvious munging. This method involves employing ASCII character codes. ASCII character codes are similar to machine language for symbolizing characters on a web page. As an example, you can represent an @ using either the @ character itself, or instead use it's ASCII character code: **@** (ampersand number-sign six four semi-colons).

Summary and Conclusion

Not surprisingly, bot technology continues to race ahead of legal and ethical concerns. As many of the countermeasures only slow down bot activity, management will not only need to develop new countermeasures but also address the legal and ethical issues that continue to arise.

It is essential, however, to keep in mind that bots are not always bad players and they have many positive uses in both commercial and non-commercial applications. In many instances, concerns center on ethical questions rather than legal issues. One of the most critical technology challenges is to establish legal and ethical boundaries for emerging technologies while developing effective bot counter-measures to mitigate commercial fraud, eliminate social engineering, and protect individual privacy of online communications.

References

Abramchuk, M. (2017, March 18). How much do chatbots cost? Quora.com. Retrieved on 03/31/2018 from https://www.quora.com/How-much-do-fully-managed-chat-bots-cost

Attkisson, Sharyl (2017). The Smear: How Shady Political Operatives and Fake News Control What You See, What You Think, and How You Vote. New York: Harper-Collins Publishers.

Bhave, A. & Budish, E. (2017). Primary-Market auctions for event tickets: Eliminating the rents on 'Bob the Broker'? (No. w23770). National Bureau of Economic Research.

Bonderud, D. (2017, June 2). Chatbots: the costs of robotic relationship building. ForbesBrandVoice. Retrieved 03/31/2018 from https://www.forbes.com/sites/adp/2017/06/02/chatbots-the-costs-of-robotic-relationship-building/#78b40e06a798

Bot definition. Retrieved 02/10/2018 from https://pc.net/glossary/definition/bot.

Cui, Y., I. Duenyas, and O. Sahin. (2014). Should event organizers prevent resale of tickets? *Management Science 60 (9)*: 2160–2179.

Douglas, N. (2018, April 2). This Twitter Bot will interview you like you'

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The Rise of Social Bots. *Communications of The ACM. 59(7),* 96-104. DOI:10.1145/2818717.

Gilchrist, K. (2017, May 9). Chatbots expected to cut business costs $8 billion by 2022. CNBC FinTech. Retrieved 03/31/2018 from https://www.cnbc.com/2017/05/09/chatbots-expected-to-cut-business-costs-by-8-billion-by-2022.html

Hayati, P., Potdar, V., Talevski, A., & Chai, K. (2011, March). Characterisation of Web Spambots using Self Organizing Maps. *Computer Systems Science & Engineering, 2*, 87-96.

How to Avoid Spambots by Project Honeypot. Retrieved 02/16/2018 at https://www.projecthoneypot.org/how_to_avoid_spambots.php

Independent Review of Consumer Protection Measures concerning Online Secondary Ticketing Facilities Presented to Parliament pursuant to section 94(3) of the Consumer Rights Act 2015 May 2016

Liu, L., Chen, S., Yan, G., Zhang, Z. (2014). BotTracer: execution-Based Bot-Like Malware Detection. Information Security, ISC 2008 Lecture Notes in Computer Science, 5222, Springer, Berlin, Heidelberg.

Neal, A., Kowwenhoven, S. & SA, O. (2015). *Quantifying online advertising fraud: Ad-click bots vs. humans.* Tech. rep., Oxford Bio Chronometrics.

New York Times. *Mr. Dickens' Readings – Sale of Tickets for the Second Course*, December 1867.

Salge, d. & Berente, N. (2017). Computing Ethics Is That Social Bot Behaving Unethically? A procedure for reflection and discourse on the behavior of bots in the context of law, deception, and societal norms. *Communications Of The ACM. 60(9),* 29-31. DOI:10.1145/3126492.

Shahmanzari, M., & Ozkah, S. (2014). Assessing the Effect of E-commerce Intelligent Bots on Online Consumers' Post-adoption Behavior for Future Use. *American Academic & Scholarly Research Journal, 6(*4), 163.