Relevance of Bots in Software and Their Impacts on Software Security

Elson Kurian and Sherwin Varghese, Member, IAENG

Abstract—Robots or bots are the terms that are more frequently heard of in the Software Industry and the academia. The purpose of developing robots was to assist humans. With the dawn of the first industrial revolution, robots became prevalent and helped in performing mundane manual labor. Now, with advancements in technology and computer science, bots are an integral part of the computer software. They have become a common user interface for many software services. Bots are preferred over humans in many services due to its perceivable passionate personality. This paper discusses about the different types of bots and their usage. For any software application, security is of paramount importance. In addition, this paper attempts to throw some light on how bots have impacted the Software Security and what steps can be taken to ensure security while using bots in applications. Bots are becoming increasingly popular. However, the number of security vulnerabilities and attacks that bots are exposed to is greatly increasing. Apart from using Bots to provide services to the user, bots are also used to spread malware and hijack distributed computer networks, compromising critical information. The paper also proposes some measures to enhance the security of bots.

Index Terms—Artificial Intelligence, Bot, Computer Security, Internet, Machine Learning and Robotic Process Automation.

I. INTRODUCTION

Bots have evolved to be the most usual communication mechanism in web applications and software services. Robots were initially developed as computer devices that assisted humans in performing manual labor work. The advancement in computer hardware led to the development of different types of robotic hardware. Industrial grade robots were developed to handle ubiquitous but dangerous tasks at production plants. Due to the vast varieties of robots available for specific tasks, the diversity in the computer hardware enforces specific robotic software to run them. Thus, robotic control software may need to be modified to support the specific hardware configuration used by the robot. With the evolution of computer science, the focus on robotics have shifted from a pure hardware-based device to software implementations. Software applications help in achieving hardware-controlled tasks, thus making cost-effective robotic devices. [1]

Currently, computer robots can be run independently of the underlying hardware and such applications that carry out automated tasks are called bots. Bots have evolved to be the most usual communication mechanism in web applications and software services. In other words, an Internet bot is a software application that executes automated tasks remotely over the Internet. Typically, bots are designed to perform mundane, repetitive tasks and are preferred as they execute these tasks at a faster rate than humans can achieve. However, improvements in automation techniques have proven that bots are very much capable of handling complex tasks which may include fulfilling business processes using Natural Language Processing (NLP). Bots constitute more than 50% of the web traffic across the Internet. There are several kinds of bots: web crawlers (web spidering) – where web pages are fetched and analyzed systematically from web servers by automated scripts, chat-bots – that make conversations disguising themselves as human beings and gaming bots that participate in network-based games. [2]

Bots assist the applications to fetch or share information, extract and analyze data, detect and monitor events and business processes. They are rapidly becoming the standard interface for interacting with software services. But a major factor is the prevalence of Big data, along with machine-learning algorithms for analyzing data across domains. Bots provide a convenient way for developers to generate immersive user experience for interacting with these algorithms and data. They bring in value in terms of integrating users with services and communication channels; thus, leading to their rapid development by software companies.

There is a huge surge in the usage and dependence of bots within computer applications. Bot frameworks allow software developers to make bots that cater to specific domain which includes social media, interactive voice response (IVR) systems, business processes and management of user experience. The main intention for building bots is to provide a seamless interface apart from the traditional UI to cater to the user’s specific needs. [4]. On the flip side, bots can cause serious threats to the internet security and they can be used to revoke user access to services (like in the case of Distributed Denial of Service – DDOS attacks); perform click fraud (CF) scamming through a network of malicious bots known as BotNets. Hence, detecting such mischievous bots and botnets is an important security concern for many organizations. [5]

II. RELATED WORKS

Bots are gaining popularity because of the improved throughput and availability considerations. Bots are developed with the main intention of automating repetitive, complex tasks which includes those within the software development processes, making the code more performant and less error prone. Bots perform tasks at faster rates, simplify tasks of the developers by allowing data replication...
across multiple business processes or by providing tools that improve the development processes. [14]

There is an intuition that the smarter the bots are, the better would the user experience be, but it’s not so. In some cases, bots have led to the degradation of the user experience by forcing users to have voice-based commands or gesture-based inputs, which are not necessary for performing the task at hand. Bots can have natural language processing capabilities and speech capabilities provided they solve the task that they are designed to perform. Business process bots, for example, may not have NLP capabilities and it does not affect the user experience. The user experience of a bot depends upon the tasks assigned to it and not on the intelligence added to it. [7] Bots can be viewed as software equivalents of robots; currently bot developments are very distant from reaching the Artificial General Intelligence (AGI) or Artificial Super Intelligence (ASI). The tasks that bot can attempt to execute, let alone the efficiency would vary greatly between the business domains. Comparing AI to a human brain, the human brain has a large network of about 100 billion neurons and around 100 trillion synapses. There is no computer till date that can match the brain. [8] The closest we have reached is to map the neurons within a mouse’s brain with the SpiNNaker, the million-core supercomputer. [9]

Out of the total Internet bots, more than half of the bot traffic are being used for malicious activities, causing troubles for enterprises having online websites. DDoS is a major form of attack targeted at a website by a botnet – network of interconnected devices hijacked by bots. DDoS attacks can cause website crashes, downtimes resulting in short term loss of businesses and they may even result in the loss of data or cause illegal access to the private organizational data. [5]

Intelligence in bots can be categorized into 3 levels. Level 1 bots are those systems that ‘remembers’ user information and user preferences. These bots rely on the information obtained from the user through interfaces which may include voice, UIs, chat or sensors. Level 2 bots are those bots that ‘learn’. These bots not only retrieve data based on the context of the user and their preferences, but also have capabilities that allow them to identify unexpected, new usage patterns, derive conclusions and perform analytics on the data. Level 3 bots are those bots that ‘understands’ the user. These bots monitor the user and proactively perform tasks on behalf of the user, thus functioning as an intelligent assistant. Few examples of level 3 bots include Google Assistant and Amazon Alexa. [11]

III. Bot Categories

Bots can be categorized in terms of their intelligence and their ability to respond to unexpected events. The advantages of bots include – (i) Adaption: bots can be designed to be context aware and would change the way they interact with users having different interests. (ii) Reasoning: bots can be designed to follow imperative first order logic rules and may use advanced ML models and AI algorithms to drive their decision making. (iii) Autonomy: Most of the bots are designed to execute tasks on its own with minimal to no user intervention; some bots may also use human inputs and learn how to make decisions with human assistance in case of events that are unlooked for. [3] In this section, we categorize the various types of bots.

A. Chat bots

Chat bot is an Artificial Intelligence software that can simulate a conversation with a user in natural language mostly through message applications, websites, mobile applications. A chatbot is one of the most advanced Question Answering system, promising expressions of interaction between humans and machines.

B. Task bots (Robotic Process Automation – RPA)

Task bots can be categorized within the Robotic Process Automation (RPA) bots. The task bots are also used in business processes to automate rule-based, repetitive task, in areas like document administration, HR, claims management, IT services such as inventory management and procuring. Task bots improves the productivity, reduces errors, and helps in saving costs. They replicate rule-based tasks and process structured data.

C. Informational bots

Informational Bots are the bots that deliver news and weather information by pushing personalized notifications by understanding user’s preferences, browsing patterns and search history. They might draw data from other APIs and provide information based on the user’s needs. [12]

D. Gaming bots

In video games, bots are AI systems that plays the game in place of a human or an opponent. Usually all the Non-Player Characters (NPC) in video games are controlled by bots; which include the opponent characters that defeat the player in the game. Having intelligent bots either as competition or as partners will help the user to have a multiplayer experience without being online. [6] Gaming bots are dynamic or static in nature. Static bots are designed to perform a set of predefined, designated tasks within the game play. Dynamic bots learn their environment as they play and behave by understanding opponents’ moves. [24]

E. Trading bots

Trading bots use algorithms to check for market deviations, detect trends and determine when trades should be made. These types of software are commonly used within forex, equity and commodities markets for several years and slowly they are moving into the world of cryptocurrency. Trading bots have invaded the share markets, especially in the 24-hour cryptocurrency domain. Wash trading, a practice involving bad bots buying and selling orders placed simultaneously to distort the true activity levels in the market, have impacted the share markets. According to the study performed by Bitwise Asset Management, as much as 95% of the unregulated exchanges are either non-economic or are fake in nature.

F. BotNets

BotNets are networks formed with a number of machines infected by malware called bots. A botnet is a network that consists of machines connected to the internet which are compromised by a malware. Botnets pose a serious threat to internet security, as they are designed to perform a wide range of malicious activities. Therefore, malicious bots and botnets need to be detected and counter measures need to be taken against them, which is an important concern for many
G. Voice bots

Voice bots are capable of verbal communication with humans, generally without another interface. They combine utility, informational, conversation, and entertainment bots. With one simple voice command items can be added to the grocery list, meeting can be scheduled with a friend, the current weather, sports news can be obtained, or it can tell a joke. Voice bots enable bidirectional context aware conversations to happen naturally. Voice bots integrate technologies from other bots as well to simulate one of the biggest milestones of Artificial Intelligence – a human conversation.

H. Bots in Software Development

1) Code bots

Code bots are those tools that help software developers to write a most portions of their software, synchronizing tasks between multiple workflows. Developers also use bots that help in peer code reviews, allowing them to automatically create Pull Requests. CROKAGE is a tool developed at Stack Overflow that provides a powerful way to search for code problems in Stack Overflow and it also provides the most relevant code snippets to descriptive coding questions using NLP.

2) Test bots

Bots play a very important role in testing. Bots like the Frued bot provided by Atlassian runs the static code analysis tools like FindBugs, CheckStyle and PMD. In addition, if there is an issue detected, then Frued raises a pull request with the suggested code change thus saving time for code quality analysis and fixes for the same. The Compare Bot, also developed by Atlassian, indicates potential User Interface bugs by comparing changes in screenshots of the application.

3) Dev-Ops bots

DevOps bots are used to rapidly improve code deployment, address the feedback delays between developers, infrastructure and operations personnel. PagerDuty is a tool that creates issues automatically, when a service fails, notifying the right people, reducing the communication overload. Pagerbot developed by Stripe, is a bot built on top of PagerDuty that enables team members to track and coordinate the PagerDuty on call schedules and efforts to respond to the issues.

4) Support bots

Support bots help mitigate the gap between developers and users. The main challenge faced is the extensive number of issues that are raised by users, most of which are for simple configurations of the software. Bots help in automating these workflows thereby reducing the number of tickets being sent to the developers; bots also answer the frequently asked questions by referring to the documentation resources and knowledge base.

5) Documenting bots

Documentation is always a challenge for developers. Atlassian uses bots to author release notes by curating comments, descriptions and code from commits and issues in the code repository. The main advantage is that these bots allow developers to write more accurate, descriptive documentation with less effort. Translation bots also assist in generating documentation for multiple languages.[14]

IV. RELEVANCE OF BOTS AND THE NEED FOR SECURITY MECHANISMS IN BOTS

Bots have gained popularity in software development during the recent years. The Gartner 2018 Technology Roadmap Survey, which included 452 service leaders from different industries, locations, contact center sizes and business models received views about 45 emerging and established technologies. It states that 25% of the customer service operations would incorporate and use Virtual Customer Assistants (VCAs) or chat bots throughout the engagement channels by 2020, as opposed to 2% in 2017.

Figure 1 shows different Business Industries and the share of respondents where chat bots are prevalent. It shows that chat bots are increasingly being used in support, customer services and after sales services; mostly replacing humans with bots to handle customer grievances and to provide solutions to customers.

Based on the dataset obtained from Google Trends, the usage and popularity of bots have been steadily increasing, [15] especially after the year 2016, as shown in Figure 2.

The Gartner research named ‘Future of Work Scenarios 2035 - Bots Go Bad’ describes how bots have a negative impact on people. Bots that attack and compromises systems in a network, spread malware and use the compromised machines to form a network of bad bots are known as BotNets. There are mainly 3 stages in the life cycle of a bot, namely – conception, spreading and attack. During the conception phase, the botnet is designed to investigate loopholes and vulnerabilities within the system. It is during the attack phase that the botnet actually begins attacking the system. Then the bot tries to replicate itself across the network of the Distributed Computing Environment during the spreading phase. During the spreading phase, the bot searches for other host machines in the network that it can replicate itself to. Once the systems are identified, the bot starts attacking the hosts and spreading further. Regardless of the topology, the model and the protocol being used, the bot installed on the infected host starts communicating with other bots via receiving and executing the commands. During the attack stage, the bot frequently communicates
online fraud, spams, clickjacking, brute force network attacks, downtimes and competitive data mining. The following are the key findings of the analysis –

1) Bad bots contributed to 21.8% of the entire web traffic in 2017; which has increased by 9.5% as compared to the year before. On the other hand, good bots contributed to around 20.4% of the website traffic; increasing by 8.7% compared to the year before.

2) 83.2% percent of bad bots impersonate themselves as web browsers, identifying themselves as Chrome or Internet Explorer or Firefox or Safari. Around 10.4% claim to be mobile browsers such as Android browser, Opera or Safari Mobile.

3) In 2017, around 82.7% of the traffic from bad bots originated from data centers. In 2016, the traffic had been 60.1%. Lower cost and high availability of cloud computing explains why bad bots are prevalent in data centers.

4) On an average website, 2-3 times account takeover attacks occur, but immediately after a breach, they are 3x more frequent, as bot operators know that credentials can be reused by users across multiple websites. [16]

with other similar bots present in the botnet. The frequency of communication increases rapidly in the attack phase than in the spreading phase. Once infected, the machine becomes part of a botnet – a network of infected computers can be by a cybercriminal remotely to carry out the illegal activities. So not only is the computer infected and the internet security compromised, but the system resources and the bandwidth are being utilized to attack other legitimate businesses or even unsuspecting users. This extreme potential for cybercrime makes botnets one of the most dangerous threats in the internet.

Networks composed of several hundreds or thousands of such infected bots have the potential to carry out malicious activities such as -

1) Create and deliver spam messages that flood millions of mailboxes in a few seconds. [10]
2) DDoS and DoS attacks, crashing entire websites and causing severe financial losses for businesses.
3) Cracking passwords and secure information through brute force, dictionary attacks.
4) Internet theft and identity theft, collecting sensitive, personal information from infected users. [13]

According to Distil Networks that presented insights by analyzing bad bot requests within the application layer and impact of automated threats, even though the awareness is high, the bot traffic generated by bad bots continue to grow at an alarming rate. [25] In addition, no industry is immune to automated botnet attacks and constant vigilance is needed to keep systems safe. Bad bots can be created by fraudsters or competitors or hackers and they play a major role in online fraud, spams, clickjacking, brute force network attacks, downtimes and competitive data mining. The following are the key findings of the analysis –

1) Bad bots contributed to 21.8% of the entire web traffic in 2017; which has increased by 9.5% as compared to the year before. On the other hand, good bots contributed to around 20.4% of the website traffic; increasing by 8.7% compared to the year before.

2) 83.2% percent of bad bots impersonate themselves as web browsers, identifying themselves as Chrome or Internet Explorer or Firefox or Safari. Around 10.4% claim to be mobile browsers such as Android browser, Opera or Safari Mobile.

3) In 2017, around 82.7% of the traffic from bad bots originated from data centers. In 2016, the traffic had been 60.1%. Lower cost and high availability of cloud computing explains why bad bots are prevalent in data centers.

4) On an average website, 2-3 times account takeover attacks occur, but immediately after a breach, they are 3x more frequent, as bot operators know that credentials can be reused by users across multiple websites. [16]
V. IMPROVE SOFTWARE SECURITY IN BOTS

Security is the most commonly overlooked aspect in an application and when uncovered, it costs more loss than what would have taken to fix the vulnerability initially. Securing the software application that uses bots can be further divided into 2 subcategories: namely, securing the bots and securing software from other BotNet attacks. This section mentions the measures that can be taken to improve the application security and make them tolerant to BotNet attacks.

A. Enhance Bot Security

The following measures would help improve software security of bots.

1) Two Factor Authentication

Two factor authentications would enforce verification of the user through 2 different communication channels. The most common form of authentication is through e-mail, which can easily be forged. Hence, an additional channel for authentication like the mobile phone One Time Passwords (OTP) can reduce the possibility of attackers maliciously creating fake accounts in the system. Though it may be the most basic step, but two factor authentications have been tried and tested to make the application more secure and is used extensively in finance, banking applications where security is a high priority.

2) End to End Encryption

End to End Encryption ensures that the messages that are transmitted between the sender and the receiver are encrypted, hence no third party can have direct access to the information, even if the transmitted data gets intercepted. This mechanism has been recently implemented in Social media platforms as well. This is one of the robust mechanisms that can prevent DDoS attacks on the application servers. [22]

3) User Identity Authentication

User Authentication helps to ensure that the web / user traffic is not generated from unknown or unauthorized sources. User ID based authentication is the most basic and common authentication mechanisms. It also helps to monitor the services provided to the user; thus, helping to adopt the pay as you use revenue model. Typical User authentication consists of a basic user id and password that needs to be provided by the user. Authentication can also be provided to the user by means of authentication tokens available to the user. [19]

4) Intent Level Authorization

Bots might need to store the context information while obtaining a request from the user. Intent based authentication comprises of 2 components – state and the context. State refers to the chat history and context is the outcome of the analysis performed on user inputs. When the intent of the user search may involve fetching the data with a different level of authorization at the backend. [23] Intent level authorization ensures that the authorized users only have access to sensitive information.

5) Channel Authorization

Bots can have multiple channels for communicating with the user. This channel can include Skype, Facebook, Slack etc. Based on the channel selected by the user, authorizations are enforced so that the user communicates to the chatbot only through the designated channel.

6) Intent Level Privacy

In compliance to the General Data Protection Regulation (GDPR) laws, organizations are forced to preserve the privacy of the user’s personal information. In addition, the system should also provide transparency on the usage of the user information. Thus, bots may be designed to have an additional level of privacy to secure critical user data. Critical data must never be revealed, even if an attack tries to compromise the backend servers. Intent of the user may be logged for auditing purposes.

7) Authentication Timeouts and Session Timeouts

Security can be further enhanced by monitoring user inactivity and causing the user session to automatically timeout, thus causing the user to login again. Access can be provided by authentication tokens that have an expiration period. Once the expiration period is lapsed, the bot access is revoked automatically. A polite way to prevent session timeouts is to ask the customer to extend the current session just before the timeout. This mechanism helps to ensure that multiple simultaneous logins of the user may not be present at any time and prevents unwanted access to hackers. [17]

8) Biometric Authentication

Bots that make secure digital transactions can be secured with additional biometric authentication mechanisms. It involves receiving biological inputs to authenticate a user. The most common biometric devices are fingerprint and iris scanners. This method is popular due to its effectiveness in ensuring the user security.

9) Self-Destructing Messages

Sensitive user information can be sent as self-destructing messages that delete the transaction or conversation made between the bot and the user. Bots can also be programmed to delete sensitive information after a timeout period. [18]

B. Identify and prevent BotNet attacks

BotNets are hard to identify and most of it would be identified only after the destruction inflicted by them is complete. The following mechanisms help Identify BotNets and help prevent BotNet attacks to an application.

1) Identify BotNet attacks

There are 3 basic mechanisms to detect BotNet attacks.

a) Static Approach

It is one of the fastest methods for identifying a BotNet attack by using a static analysis tool. Analyzing the web request, the header information and correlating it to the actual result helps to passively identify and mitigate bot attacks.

b) Challenge Based Approach

It is an improved mechanism for addressing BotNets. Using this approach (also called support-based approach), the proactive components present on the website / application would measure how a visitor interacts with the bot. The most common form of challenge-based approach is the blurry images like CAPTCHA, that cannot be bypassed easily. [21]

c) Behavioral Approach

The activity associated with a bot is carefully analyzed to ensure that it is not a Bot Net. [20] This approach helps to identify if the bot is actually achieving the tasks that it claims to perform. Most of the bots will link themselves to the parent program from which it is being invoked. If the bot behaves abnormally or if its characteristics vary from the parent program, it can be suspected to be a botnet.
d) Combined Multilateral Approach

The most effective method to identify and mitigate Botnets is through a specialized tool in a combined multilateral approach. All the three approaches are combined to identify a BotNet.

2) Prevent BotNet attacks

The following 4 measures would help in preventing BotNet attacks in the Software application.

a) Installing 'Captcha' and Challenge solvers

Captcha and other challenges in cohesion with the authentication would help mitigate botnet attacks. Thus, installing a captcha prevents the botnet from spamming pages and contact forms/

b) Install and activate security software

Install software security on the Operating system running the application. Software like anti-virus, anti-adware and firewall help in preventing unwanted requests from reaching the system. The security software needs to be updated regularly.

c) Activate anti-spam protection provided by the email host service. Ensure that strict filtering rules have been applied and update the filter if a junk email gets through the mailbox.

d) Prevent external access to the computer through ports such as file transfer protocols (FTP), Secure Shell Host (SSH), by turning off these ports when not in use.

VI. CONCLUSION

Bots are complex computer software applications that can perform a set of designated tasks repeatedly or to interact with users. Bots are beneficial as they provide results greater than human capabilities where manual work needs to be performed. They are proven to produce better user experience within software applications and assist users in navigating through application UIs. Nevertheless, bots can also be used to cause harm by attacking computer systems and affecting them over a network. Also, bots can cause a security risk within the application. If the bot is compromised, then the data accessed by the bots become vulnerable. Bots can be considered as a necessary evil in the present software domain. As with any technology, with the increase in usage, security standards and protocols need to be developed to make bots more secure and reliable; some of which are explained in this paper.

REFERENCES


