Crime Identification Using Traffic Analysis of HTTP Botnet

5

Ciza Thomas
https://orcid.org/0000-0002-1030-3000
Directorate of Technical Education, India

INTRODUCTION

A botnet is a network of malware infected systems that are controlled by an attacker through a Command and Control (C&C) channel. The attacker, also called the botmaster, controls the infected systems that are called bots or zombies. Various types of cyber-crimes are done by the botmaster with the help of these bots. A group of bots under the control of a botmaster is called a botnet. The general layout of a botnet system is shown in figure 1. The attackers use botnets to create disruption on the network or on a victim host either by utilizing the entire bandwidth in that network with bogus connections or by 100% CPU utilization on the victim host. This is through commanding the compromised bots to overload the resources of the victim machine/ network, to the point that it stops functioning resulting in denial of access. Such an attack is called a denial of service (DoS). Botnets can be used by botmaster to perform distributed denial-of-service (DDoS) attack, steal data, send spam, and access the device and its connection. These cyber-crimes are constantly evolving and hence the list of cyber threats can at no stage be considered exhaustive. Thus botnets are a great threat on the Internet by serving as the basic infrastructure for various distributed attacks.

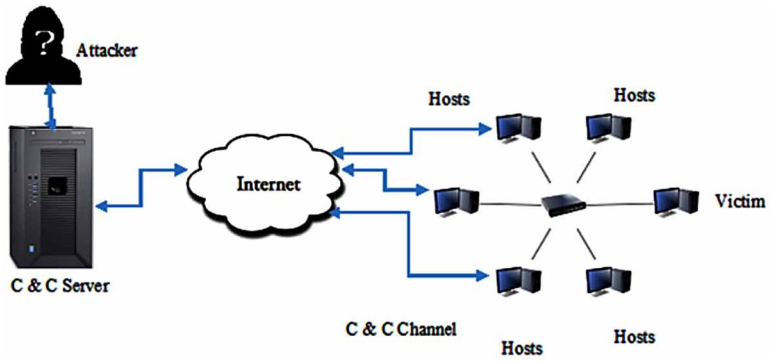
Botmasters can use HTTP protocol for the C&C channel as majority of the Internet traffic uses HTTP and hence are allowed in most of the networks. Effectively, bots hide their communication within the normal HTTP traffic as it is not easy to block this service as a precautionary measure. This fact makes the HTTP-based C&C communication stealthier. Centralised C&C channels are prone to single point of failure as the C&C channel if detected and stopped, results in the loss of the communication channel between the compromised hosts. The advantage of centralised C&C channels is that they are simple and easy to setup as highlighted in Gu and Perdisci (2008). Botmasters have moved to peer-to-peer (P2P) C&C architecture to make their bots more powerful and stealthy.

Bots run automated programs that are designed to execute specific scheduled activities or to respond to commands in a particular manner. Hence, it is expected that the botnet generated traffic should always be having an apparent structure and regularity in the behavioral pattern. This is because the normal user behaviour is totally unpredictable, random and complex. This is attributed to the innumerable online applications and resources available for users. Hence the normal traffic differs from the botnet communication traffic, which is systematic and consistent in behaviour.

Several detection strategies have been developed in the available literature for botnet detection like up-to-date anti-virus software, signature-based intrusion detection systems for IRC/botnet traffic and traffic flow monitoring for known C&Cs. These detection techniques differ based on the C&C mechanism being centralised architecture (IRC, HTTP) or peer to peer (P2P) architecture or hybrid P2P/Centralised architecture. Detection also varies depending on other factors such as area of deployment, data captured for the detection system, etc.

DOI: 10.4018/978-1-5225-9715-5.ch074

Figure 1. Typical botnet system



This work proposes a technique to collect and analyse HTTP botnets. The deficiency of publicly available botnet datasets creates the characterization of botnet traffic difficult. Hence, HTTP botnet analysis plays a key role in the design of an effective detection system due to the fact that the foremost task in the process of mitigation of a threat is its characterisation. The network traffic generated by the HTTP botnets is analysed based on various features which could be used to develop an effective detection model. In this work a framework was developed in order to build HTTP botnets in a controlled environment. Analysis of the botnet traffic shows that periodicity is a main feature of HTTP botnets. This is because of the HTTP bots periodically contact the botmaster for commands and control messages by connecting to particular URLs or web pages. In addition the bots also report their status and attack results to the botmaster. This can be utilised to build behavioural detection models at network layer as reported in the work of AsSadhan et al. (2009). Signatures of the bots that were set up are also obtained, which can be used in signature-based detection. Further analysis was done using machine learning based classification as well as periodicity analysis. The results demonstrate the superior detection performance with 100% accuracy and detection of the proposed method using the hybrid periodicity analysis.

The rest of this paper is organized as follows: The state-of-the-art of botnet structures and its detection approaches are discussed in the next section. The system architecture is subsequently discussed followed by solutions and recommendations. The paper concludes after identifying the future directions of research.

BACKGROUND

A botnet is usually established by a botnet writer developing a program, called a bot or agent, and installing the program on compromised computers on the Internet using various techniques as demonstrated by Yu et al. (2015). As botnet has now become one of the major threats in the present attack landscape, many researchers are rigorously exploring the detection, mitigation and prevention of botnets. Available literature provides details of previous research work that are carried out aimed at distinguishing or detecting HTTP-based bots, many of which use network communication features as identifiers of botnet behaviour. In the work of Khattaket al.(2014), the authors have given a detailed review of botnet behaviour, detection and defence. Several studies and researches have been carried out in order to collect and analyse the malware activities like in the work of Baecher et al.(2006), Cooke et al.(2005), and Freiling (2005). A detailed study of botnet activities by a multifaceted approach to collect malware is carried out in the work of Rajab et al. (2006). Detection systems have been developed for different