Yadigar N. Imamverdiyev[1], Gulnara B. Garayeva[1,2] DOI: 10.25045/jpit.v08.i1.11
[1]Institute of Information Technology of ANAS, Baku, Azerbaijan
[2]Sheki branch of ASPU, Shaki, Azerbaijan
[1]yadigar@lan.ab.az, [2]qarayevagulnare@mail.ru

BOTNETS AND METHODS OF THEIR DETECTION

A botnet is a network of infected computers and the C&C servers, which are controlled by botmasters. Botnets play an important role in cyber attack infrastructure, sometimes millions of computers are involved in these networks. Botnets are constantly evolving, their structure, used protocols, infection methods, purposes of attacks are constantly changing. The paper studies the architecture of botnets, classification of botnets according to various criteria and botnet detection methods.

Keywords: *botnet, C&C server, honeypot, DDoS attack, botnet detection method.*

Introduction

Botnet – the combination of the words "robot" and "network" is a special network of computers infected with bot malware. Bot is a stealthily installed malicious software (malware) performing certain actions at user's computer using the resources of the infected computer without the user's permission. Botnets are often called "zombie computer networks" which automatically realize given commands.

Botnets are used in many online cyber-crimes, large-scale DDoS attacks (Distributed Denial-Of-Service), spam distribution, click fraud, confidential information theft and other malicious activities. The number of computers infected by botnets and their geographical distribution depend on botmasters' program codes, which can be easily accessible, and a wide range of supporting communications used by them. Such networks are mainly used for profit. Hence, botmasters can rent botnets at any time and realize online attacks. It makes it quite difficult to detect cybercrime. Botnets involving millions of computers are known and new ones are emerging every day.

In the past decade, numerous scientific achievements have been made in the field of developing botnet detection methods. Botnets have been approached from different aspects and detection methods based on different technical principles have been developed. Most notable among these approaches and significantly successful results are based on Machine Learning (ML).

The increase in the Internet services and their indispensable role in our lives has caused a series of problems such as user data security, privacy, and integrity. Over the past 20 years, the Internet and Internet-based applications have been complementing our lives. The implementation of education, health care, banking, public and social life and other areas over the Internet has complicated the solutions of security issues. The basis of Internet security threats is malware. The user data in the computer infected with malware also goes under the risk. In parallel with the methods of combating the malware, the spread and infection of such software mechanisms, and the types of malicious activity have also changed and developed.

Bot malware [1, 2] easily impose computer programs like other malware- viruses, Trojans, rootkits, worms and others. However, the main difference and advantage of bot malware is its capability to keep in touch with malicious using the C&C (Command & Control) channel. Botmasters remotely control bot malware via this channel and realize it goals more easily.

Botmaster develops high level distributed bot platforms for large scale, malicious and illegal activities that can cooperate with each other. Botnets can cover PCs, corporate and educational networks, control systems and so on. The power of botnets is measured not only with their number, but also with their harmful activities. Such networks send spam, generate DDoS attacks, realize click fraud, spread malware, illegal content and advertising and collect confidential information, and attack the critical infrastructure [3].

Botnet neutralizing includes detecting botnets and taking appropriate protection measures. There are different approaches to the botnet detection, which are mainly based on monitoring of

network traffic generated by botnets and the analysis of bot behavior. ML methods are most frequently used methods in this field, which detect malicious traffic according to the generated characters. Approach to ML-based detection is to ensure the generation of appropriate characters from the selected network traffic and their effective detection.

Struggle against Botnet can be conducted in two phases. First, preventing the boat infection (direct responsibility of the user), and second, detection of infected devices and termination of botnet performance. The variety of infection methods used by botnets, insufficient exploration of their architecture, and unawareness of the users about the ways of protecting are the main reasons for the botnet development.

At present, botnets and the areas of their use are rapidly expanding. Botnets establishment and management techniques, as well as the types of attacks are increasing, and accordingly, the development of new control methods still remains urgent.

Brief description of botnets

Malware botnet is a network of infected computers that can be controlled by an attacker. Thus, each device on the network is called the bot, and malicious attacker that controls botnets - botmaster. Botmaster uses the botnets for various purposes. Botmasters start their activity by collecting bots, i.e., by occupying new computers. For this, various methods are used, such as viruses or other malware, e-mail, spam, attractive images with different purposes, advertising and so on. The most commonly used modern methods of bot attracting are the methods realized by using the potential of social networks with a large number of users (spreading fake accounts, pages, advertising, etc.). Users upload such means, spread by the botmasters, as a result, the software is installed on the computer. Once the computer is infected by bot it is connected to the botnet control center and waits for command. The working condition of the infected computer is enough to be controlled by the botmaster. It is used without the permission of the user.

Botnets are used for different purposes. The most common ones include [4]:

1) Spam distribution. According to some sources [5], 90% of Internet traffic is covered by spam. Thus, over 95% of spam is sent by botnets. The main advantage of spam distribution for botmasters is that the party receiving the spam is unable to legally compete with the party that sends the spam, and spreading spam in large volumes;

2) DDoS attacks. Botmasters tend to access a large number of Web sites at the same time using botnets. In this case, the web-server delay or even denial of service condition occurs. The most commonly used DDoS attack type is Syn-attacks (Syn Flood). In addition, another DDoS attack type is the Degradation of Service;

3) Information theft. It covers many areas: secret data (economic, scientific, medical, etc.), government information (military secrets, state documents, etc.), plastic card information, personal data (address, phone number, username/password, PIN-codes, etc.);

4) Click fraud. Botnets are used to artificially increase the number of users of Web-sites, personal blogs, social network pages, as well as to increase (or decrease) the number of clicks for the electronic voting system;

5) Creating the infrastructure for phishing attacks;

6) Distributing other malicious programs, such as spyware, distribution of advertising materials and so on.

Botnet phenomenon

For variety of C&C and purposes of malicious activities, the botnet is one of the phenomenons that should be comprehensively and thoroughly studied. In addition, botmasters periodically improve botnet’s working methods in order to complicate the detection or even make it possible. To better understand the phenomenon of the botnet’s lifecycle, C&C channels, the methods used for flexibility and sustainability should be studied.

Botnet Life-cycle

Life-cycle of botnet is one of the decisive factors for the selection of detection methods. In many works, the life-cycle of botnet is divided into five phases (Figure 1) [6, 7]:

- infection phase;
- repeated infection;
- communication phase;
- attack phase;
- support and update.

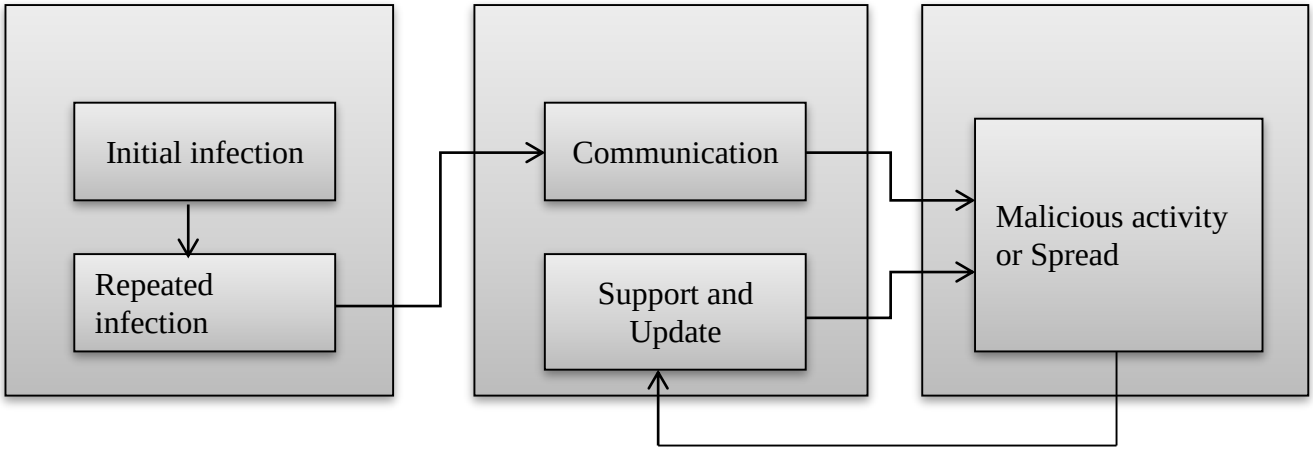

Figure 1. Life-cycle of a botnet

The first phase infects computers with the bot program in any way. The infection phase itself is divided into two sub-phases: initial and repeated infection. During the initial infection, the computer is infected with "loader". Initial loading can occur through different ways: unauthorized loading from malicious websites, e-mail message attachments, infected storage devices and so on. After the first successful loading, the repeated loading starts. Botnet malware is downloaded to a computer by using different protocols, such as FTP, HTTP, HTTPS, or any P2P protocol and waits for communication.

Communication phase includes adoption of guidelines between C&C server and bots, and their update by botmaster, as well as creates stable contact to check the status of bot. Communication is carried out through several operations: connection attempts to C&C server after successful infection, re-connection attempts after enabling the infected machine, periodical connection attempts to check the status of the infected machine, malicious code updates by C&C or attempts to ensure spread of commands to other bots. Communication between zombie computers and the C&C server is implemented in different ways.

In attack phase, the bots are directly engaged in malicious activities in accordance with the botmaster's objective. DDoS attacks, starts spam e-mail campaigns, launches theft information dissemination, and manages online systems and requests, and so on. At this stage, the bots are involved in distribution and infection of new boats.

C&C communication channel

The classification of botnets is based on the protocols used for their architecture and management. One of the main categories that characterize the botnets is C&C servers. Botmasters use such special servers to keep in touch with bots [8]. As C&C, the hidden servers, the geographic area of which is unknown, are often used. Botmaster builds C&C server to handle the botnet, and the infected computer connects to the server via any channel (e.g., IRC (Internet Relay Chat)) and waits for commands from the server. One botnet may include more than one C&C server, which makes it difficult for its detection and destruction. Botnets can be classified in several ways

according to its architecture. [9] classifies the botnets as: centralized, P2P (Peer-to-Peer) and hybrid (tree) architecture:

1) *Centralized C&C*

Centralized C&C model is one of the key models used for botnets. Most popular botnets - AgoBot, Sdbot, Rbot and others refer to a centralized model. In this model, botmaster chooses a high-speed channel to keep in touch with bots. Usually, in such model, C&C server obeys the computer to "corporate" using network protocols as IRC and HTTP. Once a new computer is infected with bot, it connects to botnet by contacting with C&C server. Once joining to the specific C&C server it waits for botmaster's command.

The use of centralized C&C model has the following advantages:

- more simple infection and "privatization" in the centralized C&C model due to the use of easily accessible programs (such as IRC scripts and IRC bots). Using such a model, the botmaster can simultaneously handle thousands of bots. When building profit botnets, centralized C&C model, which manages more bots and ensures maximum income, is chosen;
- very short delay time of messages in the centralized model, which provides botmaster to easily manage boots and organize attacks.

However, there are several disadvantages of the centralized C&C model, so that all exchanges are conducted over the centralized server, which weakens the C&C server. If the C&C server is discovered, the entire botnet is collapsed (Figure 2).

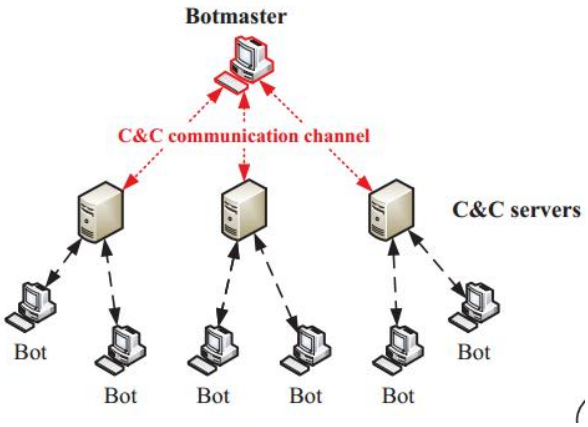

Figure 2. Centralized C&C

2) *P2P-based C&C*

P2P-based C&C model is one of the botnet models considered to be more detection resistant. Although the number of P2P-based botnets is very small, compared to the centralized model, its detection and destruction is more difficult. In non-central botnets, the bots is connected to numerous infected machines of a zombie network rather than the control center. The commands are sent from bot to bot, each bot has a list of several "neighbors' "address, and the command received from any of them is sent to another neighbor, thereby spreading the commands. In this case, botmaster should have a direct access to, at least, one of the computers which is included in the botnet to manage the botnet [10, 11].

Though, the botnets based on this model is growing and its detection is more difficult, there are also some disadvantages. First, P2P systems enable communication between very small numbers of user groups (10–50). Compared to the centralized model, this scale is very small. Second, P2P systems cannot assure the timely delivery and non-delayed distribution of messages. One or more non-active bots can weaken the network distribution enough. Therefore, such botnets are difficult to be managed. These two reasons are reducing wider acceptance and use of P2P model (Figure 3).

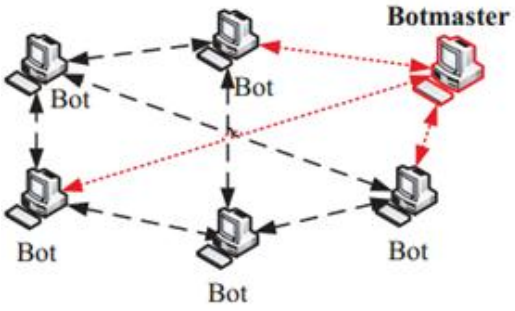

Figure 3. P2P-based C&C

3) *Hybrid C&C*

Hybrid contact is one of C&C contact forms used by modern botnets. In this case, it is used together with non-central architecture for centralization and disclosure resistance to ensure short delays. In such architecture, the bots are divided into two classes: Proxy bots and working bots. Proxy bots constantly provides the interaction between botmasters and other bots and the spread of commands, while, working bots just connect to proxy bots and wait for the commands from botmasters (Figure 4).

One of the characteristic categories for botnets is the operating mechanism of C&C servers. These mechanisms are essential for finding new boots to and making them dependent on botmasters. The most commonly used techniques are the followings [12, 13]:

Hard-coded IP address. In this method, the first infected bot contacts with C&C server with fixed IP address. The disadvantage of this method is that C&C servers using Hard-coded IP address can easily be detected and the connection can be blocked. In this case, if the connection between the server and the bots fails, the botnet is completely destroyed. Therefore, this method is almost not used for modern botnets.

DNS flux. Botnets use domain names generated with the cryptographic methods using DGA (Domain Generation Algorithm). This technology extremely complicates finding all possible C&C addresses for static systems. If the C&C server is closed by the owner, the control easily transferred to the new server. The bot disconnected from the old server sends DNS request and connects to the new C&C server. Using dynamic DNS names, the botmaster can recover C&C server when it loses its function. C&C servers are often deliberately changed in order to complicate the detection.

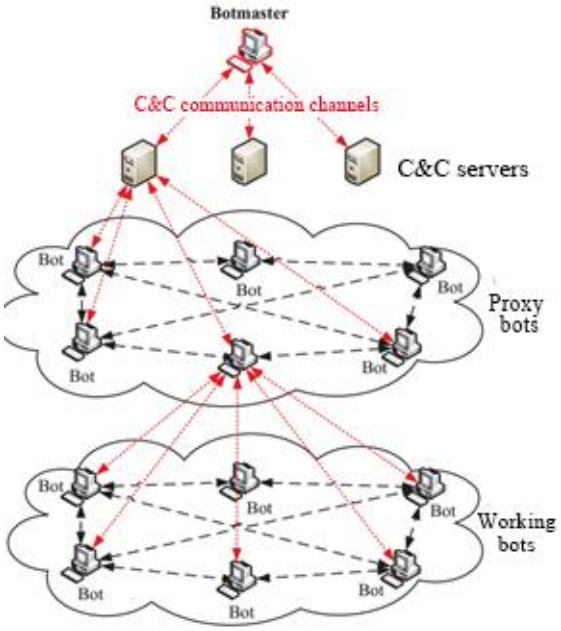

Figure 4. Hybrid C&C

IP flux. This technology is also known as FFSN (Fast Flux Service Networks). FFSN is a DNS technology used to hide malicious botnet servers in the network. The basis of this technology is maintaining more than one IP address for one domain name and constantly changing DNS cache of these addresses. Two kinds of the fast flux technology are available: single and double.

To transmit the commands of the botnet's owner to the bot computer, it is necessary to establish a network connection between the bot and the computer, which sends the command. The whole network is based on network protocols, which specify the communication guidelines of computers within the network interaction. Therefore, botnets can be classified according to the used communication protocol. Depending on the type of used network protocols, the botnets are divided into IRC, IM (Instant Messaging), web and other (TCP, ICMP, UDP, etc.) oriented groups.

Classification of botnet detection methods

Since the early 2000s, a large number of approaches based on the different technical principles for the botnet detection have been proposed. Detection approaches are classified into two main classes: host-based and network-based. Host-based methods examine the behavior of malicious bot program on the infected machine. For this purpose, user level behaviors as the application and system logs, active processes, key logs, resource usage and so on are studied.

Network-based detection often analyzes the network stream on router or firewall for the botnet detection. This class of detection methods ensures detection through the recognition of network traffic generated by bots on all three stages of bot lifecycle. Such detection methods are often called Intrusion Detection Systems (IDS) or Intrusion Prevention Systems (IPS) [14, 15].

In addition, as a result of the simultaneous application of two approaches hybrid detection methods have been developed in the work [16, 17]. This method class analyzes both behavior of the user, and also the network flow.

It should be noted that host-based methods are quite vulnerable to resistance improvement methods used at the user level. Botmaster applies various methods for bot programs not to be detected in the infected machine, furthermore, host-based methods can only detect one infected bot and its C&C communications. Network-based methods study network activity of (spam distribution, DDoS-attack, etc.) infected machines, as well as the attempts to contact any C&C server. The main advantage of network-based methods is the capability to observe and discover numerous infected machines.

Network-based detection

Network-based detection [18] techniques are based on the analysis of network flow to detect infected machines. Network traffic can be analyzed at the packet and network flow level. Network flow is usually identified by five major indications: Source and Destination IP addresses, Source and Destination ports, and protocols. However, in real detection methods, more flow indications can be analyzed. Network-based methods themselves are divided into several classes by their application points, operation confidentiality, and the basic working principles.

The variety of network-based detection methods primarily depends on which part of network traffic the method is applied. Obviously, network-based methods aimed at detecting more bots should handle a large volume of data [19].

Detection methods are divided into two active and passive classes according to the confidentiality function. Passive detection approaches do not interfere the botnet operations; they just monitor them, so that it makes detection process confidential and not noticed by the botmasters. Active detection methods can actively intervene the harmful activities of botnets or to the attempts to contact with any C&C server. Typically, such methods are aimed at heuristic C&C connections and provide high detection accuracy at the expense of flexibility and commonality of the approach. The main advantage of passive approaches, in this regard. is its potential to detect large-scale types

of botnets only according to observed malicious traffic signatures (characters). Most available detection methods are passive; there are very few active methods.

According to the key functional working principle, the detection methods are divided into two classes: *signature-based* and *anomaly-based* classes. Signature-based methods are based on the recognition of characteristic traffic images, called "signature". This class of detection techniques can be applied at any stage of the botnet lifecycle, and they provide high accuracy detection of known botnets. However, the main disadvantage of the method is its capability to detect only known threats, and this approach requires periodic update of signatures.

Anomaly-based detection methods are based on the detection of malicious traffic in accordance with the anomalies occurring in the network traffic. Such methods based on the easily detectable network events, such as the changes in the level of network traffic, delays and so on. This class of approaches can work at the packet and flow level. Anomaly-based approach applies statistical analysis, ML, graph analysis and other methods. Unlike signature-based methods, anomaly-based detection approaches can detect new threats and they are resistant to the botnets' confidentiality measures. The traffic used by the modern botnets are quite similar to the normal traffic, so one of the important issues is the adoption of events as an anomalous phenomena. In addition, as it is important to analyze very large-scale data during the anomaly detection, the most commonly used approaches are ML-based analysis methods. Because, ML methods offer automated recognition with the help of traffic images and ensure recognition of malicious traffic images when there is no information about malicious traffic features in advance.

Many detection methods with the application of ML methods have been proposed so far. The most effective results are achieved by the techniques as BotMiner [20], BotGAD [21], BotSniffer [22], BotHunter [23], EFFORT [24] and others.

ML-based botnet detection

ML [25, 26] is a branch of artificial intelligence and the main goal of which is the knowledge summarization based on finite number of previous experiences sand finding useful patterns for previously unknown events. The main advantage of ML methods is finding the needed knowledge out of large amounts of data. ML methods are applied in many areas, as statistics, artificial intelligence, information theory, cognitive science, management theory and so on. One of the application fields of ML algorithms is network traffic detection associated with bots.

Two main classes of ML algorithms are distinguished:

- supervised learning;
- unsupervised learning;

Supervised learning is one of the most widely used ML methods, where the algorithms are taught according to the input samples and the output samples matching them and then used in forecasting output for any new access. Supervised learning is used to specify input data to any identified class and to give assessment forecast at output.

Unsupervised learning issues may include finding similar samples within targeted input data - clustering, identifying data distribution in input space – logging (density estimation) or bringing the data from multi-dimensional space to 2- or 3-dimensional space for their description.

In most cases, the anomaly-based botnet detection uses supervised or unsupervised ML algorithms such as classification or clustering. Network traffic is analyzed in both directions (input and output) and the features are extracted in the package level. Extracted traffic features describe specific traffic flow or specific servers and hosts on the network. With the use of Supervised ML algorithms the scheme of botnet detection methods is as follows (Figure 5).

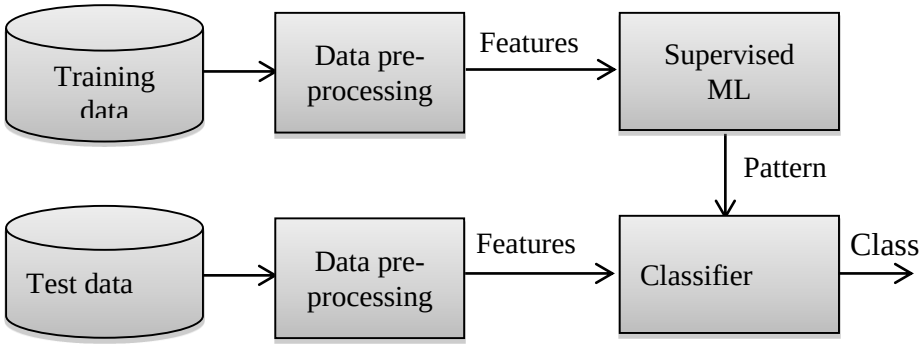

Figure 5. Supervised detection scheme

Supervised ML algorithms are first taught through the training data and the functions are established according to appropriate input and outlet. This function is then used as a model to classify the inputs of test data. To use supervised ML methods both data sets - test and training data must be initially processed. Initial processing of data includes the introduction of a number of processes and provides selection of features. Selecting the right features for the proper functioning of ML algorithm after its implementation of in practice is one of the hardest problems. Because the right or wrong selection of features directly influences the quality of detection. The features can be selected based on heuristic properties of botnets, the type of C&C communication channel, used protocol, and so on. Most commonly used ML methods for botnet detection are supervised SVM (Support Vector Machines), artificial neural networks, decision trees, Bayesian networks and others.

Unlike supervised learning, unsupervised learning techniques are used for bot-related data clustering. The main characteristic property of unsupervised ML algorithms is not requiring the training in advance. The scheme of botnet detection with the introduction of unsupervised ML algorithms is as follows (Figure 6):

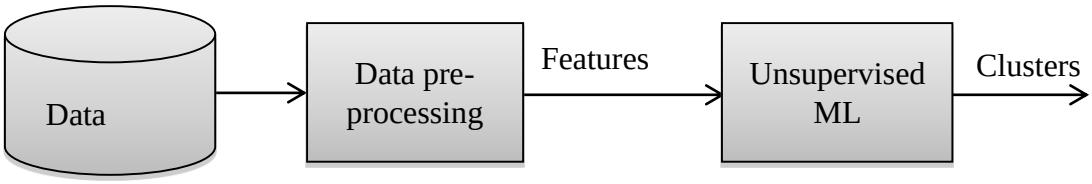

Figure 6. Unsupervised detection scheme

For extraction and selection of features in these methods the data clustering is realized using initial processing and unsupervised ML algorithms. The main problem of such type of detection methods is selection of the relevant features and naming the clusters. The most popular unsupervised ML algorithms used for botnet detection are k-means, X-means and hierarchical clustering.

When applying ML methods for botnet detection on both scenarios, initial data processing and feature extraction, which can successfully identify the botnets, are important processes. In addition, modern detection methods are often applied in combination with ML algorithms for high quality detection.

Basic principles of detection methods

ML-based botnet detection is considered to be a modern approach, and has a number of key features as follows:

1) Generality;
2) Stealthiness;

3) Timely detection;
4) High detection performances;
5) Robustness on evasion techniques.

The generality means the possibility of being applied in the detection of various types of botnets, regardless of the same distribution mechanisms, attack targets, and the used C&C channel. Detection method can be applied on one of infection, contact, or attack stages of botnet's lifecycle. Obviously, botnet detection in the initial or repeated infection stage could be more effective in terms of preventing their subsequent activities. However, this is quite difficult, since the information for direct detection is insufficient and the diversity of methods used on the infection stage hides malicious bots "well".

Detection approaches applied in connection stage are aimed at connection protocols (IRC, HTTP, P2P etc.) and botnet network topologies. The first such detection techniques were mainly aimed at IRC traffic. However, advanced botnets are using P2P and HTTP protocols more, and therefore, detection methods because are focused on these protocols.

The methods operating on attack stage are focused on the specific attack attempts (spam, DDoS, etc.). Such detection methods usually analyze activity densities of the computers infected as a group, and significant deviations from normal traffic density can be taken as a bot warning.

In addition, generality of botnet detection depends on the bot-related heuristic assumptions of the approach and on how they show themselves in real botnets. Specific botnet types or the approaches targeted at specific lifecycle, in many cases, provide more effective results than the universal methods aimed at all types of botnets. It should be noted that the malicious needs the bots for long-term use, and they do not change communication protocol or C&C communication channel unless it is necessary. However, the specific methods poorly adapt to any change in the nature of botnets.

Stealthiness– the specification of detection approach not to be noticed by the attackers, so that all passive methods are confidential during their lifecycles. Most proposed detection approaches are passive and provide stealthiness requirements in their activities.

Timely detection - often requires method to work on-line, which leads to real time analysis of large-scale data. However, it is difficult to say that many methods provide the term of timely detection.

The quality of detection

The quality of the detection methods is measured by the achieved real results and the value of deviation technology that they are sensitive to. The method is usually measured via a specially selected database. Test data can contain malicious and non-malicious traffic data. Correctly chosen data set is one of the basic conditions for the assessment of the detection quality. Malicious traffic is presented as the traffic generated by the botnets, while non-malicious traffic is often called "background" and contains "clean" traffic created by the non-malicious hosts. Bot-related traffic can be obtained with the following guidelines [27]:

1) Bot-related traffic can be collected by a special built-in honeypots;
2) Bot-related traffic can be collected from the fully controlled organization network;
3) Bot-related traffic can be collected from the partly controlled organization network.

Selecting non-malicious traffic itself is also problematic enough and can be collected in different ways, for example, through the use of static traffic generator, via sniffers in the local networks, and even from ISP (*Internet Service Provider*) networks.

Quality metrics of the detection approaches differ, the most commonly used metrics are as follows:

1) *True positives rate* (*TPR*): $TPR = \frac{TP}{TP+FN}$.
2) *True negative rate (TNR*): $TNR = \frac{TN}{TN+FP}$.

3) *False positive rate (FPR)*: $FPR = \frac{FP}{FP+TN}$.
4) *False negative rate (FNR)*: $FNR = \frac{FN}{FN+TP}$.
5) *Accuracy*: $accuracy = \frac{TP+TN}{TP+FP+TN+FN}$.
6) *Error*: $error = \frac{FP+FN}{TP+FP+TN+FN}$.
7) *Precision*: $precision = \frac{TP}{TP+FP}$.

where *TP* – denotes the number of correctly classified positive samples, *TN* - the number of correctly classified negative samples, *FP* - the number of incorrectly classified positive samples, *FN* - the number of incorrectly classified and negative samples. However, all abovementioned metrics for measuring the quality of the approach is not always used. Most commonly used metrics are *TPR* and *FRP* percentages. *TPR* percentage to be large enough and *FRP* percentage to be small enough to define the quality of the detection approach.

Conclusion

Botnets is one of the main tools in the implementation of large-scale cyber-attacks to the critical information resources of countries. Both the number of DDoS attacks, as well as other problems caused by botnets proves how dangerous they are. Botnets prevention, their timely detection and implementation of the counter-measures against them are important. The analysis of this research shows that most of approaches to the botnet detection are based on Machine Learning and Data Mining methods. Future studies are planned to be conducted in this field.

References

1. Liu J., Xiao Y., Ghaboosi K., Deng H., Zhang J. Botnet: classification, attacks, detection, tracing, and preventive measures // EURASIP Journal on Wireless Communications and Networking, 2009, pp.1–12.
2. Li C., Jiang W., Zou X. Botnet: survey and case study / Proc. of the 4th International Conference on Innovative Computing, Information and Control, 2009, pp.1184–1187.
3. McKewan A. Botnets – zombies get smarter // Network Security, 2006, vol.2006, no.6, pp.18–20.
4. Schiller C.A., BinkleyJ., Evron G., Willems C., Bradley T., Harley D., Cross M. Botnets: the killer web app. Syngress, 2007, 480 p.
5. Rodrigues N., Sousa R., Ferreira P.S., Nogueira A.M. Characterization and modeling of top spam botnets // Network Protocols and Algorithms, 2012, vol.4, no.4, pp.1–26.
6. Silva S.S., Silva R.M., Pinto R.C., Salles R.M. Botnets: A survey // Computer Networks, 2013, vol.57, no.2, pp.378–403.
7. Feily M., Shahrestani A., Ramadass S. A survey of botnet and botnet detection / Proc. of the 3rd International Conference on Emerging Security Information, Systems and Technologies, 2009, pp.268–273.
8. Zeidanloo H., Manaf A. Botnet command and control mechanisms / Proc. of the 2nd International Conference on Computer and Electrical Engineering (ICCEE'09), 2009,vol.1, pp.564–5683.
9. Trend Micro. Taxonomy of botnet threats. Technical Report, 2006. http://www.cs.ucsb.edu/ kemm/courses/cs595G/TM06.pdf
10. Rodríguez-GómezR.A., Maciá-Fernández G., García-Teodoro P., Steiner M., Balzarotti D. Resource monitoring for the detection of parasite P2P botnets // Computer Networks, 2014, vol.70, pp.302–311.
11. Singh K., Guntuku S.C., Thakur A., Hota C. Big Data Analytics framework for peer-to-peer botnet detection using Random Forests // Information Sciences, 2014, vol.278, pp.488–497.

12. Sharifnya R., Abadi M. DFBotKiller: Domain-flux botnet detection based on the history of group activities and failures in DNS traffic // Digital Investigation, 2015, vol.12, pp.15–26.
13. OpenDNS Security Whitepaper. The role of DNS in botnet command &control. http://info. opendns.com/rs/opendns/images/OpenDNS_SecurityWhitepaper DNSRoleInBotnets.pdf
14. Jabez J., Muthukumar B. Intrusion Detection System (IDS): Anomaly detection using outlier detection approach // Procedia Computer Science, 2015, vol.48, pp.338–346.
15. Kacha C., Shevade K.A. Comparison of different intrusion detection and prevention systems // International Journal of Emerging Technology and Advanced Engineering, 2012, vol.2, no.12, pp.243–245.
16. ZengY., HuX., ShinK. Detection of botnets using combined host and network level information /IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 2010, pp.291–300.
17. ZengY. On detection of current and next-generation botnets. Ph.D. thesis. The University of Michigan, January 2012.
18. Zhao D., Traore I., Sayed B., Lu W., Saad S., Ghorbani A., Garant D.Botnet detection based on traffic behavior analysis and flow intervals // Computers &Security, 2013, vol.39, part A, pp.2–16.
19. Stevanovic M., Pedersen J.M. Machine learning for identifying botnet network traffic, Aalborg Universitet, Technical Report, 2013, 29 p.
20. Gu G., Perdisci R., ZhangJ., Lee W. Botminer: Clustering analysis of network traffic for protocol- and structure-independent botnet detection / Proc. of the 17th Conference on Security Symposium, 2008, pp.139–154.
21. Choi H., Lee H. Identifying botnets by capturing group activities in DNS traffic // Journal of Computer Networks, 2011, vol.56,pp.20–33.
22. Gu G., Zhang J., Lee W. BotSniffer: detecting botnet command and control channels in network traffic / Proc. of the 15th Network and Distributed System Security Symposium (NDSS), 2008, pp.1–18.
23. Gu G., Porras P., Yegneswaran V., Fong M., Lee W. BotHunter: Detecting malware infection through IDS-driven dialog correlation / Usenix Security, 2007, vol.7, pp.1–16.
24. Shin S., Xu Z., Gu G. EFFORT: Efficient and effective bot malware detection / Proc. of the 31th Annual IEEE Conference on Computer Communications (INFOCOM'12) Mini-Conference, 2012, pp.71–80.
25. Masud M., Khan L., Thuraisingham B. Data Mining Tools for Malware Detection. Taylor & Francis Group, 2011.
26. Dua S., Du X. Data Mining and Machine Learning in Cybersecurity, CRC Press, 2011, 248 p.
27. Aviv A.J., Haeberlen A. Challenges in experimenting with botnet detection systems / Proc. of the 4th Conference on Cyber Security Experimentation and Test (CSET'11), 2011, pp.6.