A Method to Detect Internet of Things Botnets

Anton O. Prokofiev [#1], Yulia S. Smirnova[#2], Vasiliy A. Surov
Dept. of Computer Systems and Technologies
National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)
Moscow, Russia
[1]AOProkofiev@ieee.org, [2]manzanilla17@bk.ru

***Abstract*— The main security problems, typical for the Internet of Things (IoT), as well as the purpose of gaining unauthorized access to the IoT, are considered in this paper. Common characteristics of the most widespread botnets are provided. A method to detect compromised IoT devices included into a botnet is proposed. The method is based on a model of logistic regression. The article describes a developed model of logistic regression which allows to estimate the probability that a device initiating a connection is running a bot. A list of network protocols, used to gain unauthorized access to a device and to receive instructions from common and control (C&C) server, is provided too.**

Keywords—IoT; botnet detection; Internet of Things; cybersecurity

I. INTRODUCTION

An emerging trend in the field of Information and Communication Technologies (ICT) is the increasing popularity of the Internet of Things (IoT). The IoT is a dynamic global network infrastructure with self configuring capabilities based on standard and interoperable communication protocols where physical and virtual "things" have identities, physical attributes and are seamlessly integrated into the information network [1]. This is the world of interconnected "things", where humans interact with devices and devices in turn interact with each other (M2M [2]).

Unprecedented in its scale and speed distribution of devices enabled by various types of wireless technology such as Bluetooth, radio frequency identification (RFID), Wi-Fi and telephonic data services [3] resulted in the IoT becoming one of the primary trends of high technology. As early as 2010 the number of network connected devices has exceeded the world`s human population. According to prediction of Cisco researchers almost 50 billion IoT devices will have been connected to the Internet by 2020 [4]. However, a low level of information security is still remaining one of the major issue related to the IoT [5]. Failures to reset unsecure passwords, as well as a lack of protection against brute-force attacks are widespread security problems typical for a great number of IoT devices [6]. Additionally, a lot of mobile applications designed to control and monitor devices via network do not support modern security standards such as Secure Socket Layer (SSL) [7] to encrypt communications [8]. Moreover, plenty of IoT devices do not generally allow to reset default credentials or to install software updates that leads to impossibility of fixing vulnerabilities, e.g. such as a disclosed in October 2017 weakness in the encryption protocol WPA2 [9], which is widely used in the majority of modern wireless networks. Thereby, the IoT is becoming increasingly popular as a powerful tool of cybercriminals. According to Gartner analysts 25 % of cyber attacks will have involved IoT devices by 2020 [10].

IoT devices are compromised by cybercriminals in order to install ransomware programs, to steal personal information, as well as to include them into a botnet . Botnets are used in a huge variety of cybecriminal activities, among which the most popular are phishing campaigns, spamming, malware delivering, as well as Distributed Denial of Service (DDoS) attacks [11]. In the light of recent massive cybercrime, the IoT poses a serious security threat to any structure connected to the Internet. Powerful DDoS attacks on such companies as Dyn, Amazon, Twitter and Reddit were performed by the Mirai botnet, which appeared in 2016 and consistsed of approximately 500,000 compromised IoT devices [12]. Victims of botnets mainly are routers, digital video recorders (DVRs) and IP-cameras. At the same time, new inproved versions of malware used to create large-scale botnets appear continually. Therefore, an effective technique to detect devices controlled by bots is essential in order to prevent botnet attacks.

This article presents the results of a study of botnets consisting mainly of IoT devices. A method to detect IoT botnets at a propagation stage by performing brute-force attacks on targeted devices is provided.

The second section of this article provides a description of a botnet lifecycle and main characteristics of the most powerful botnets. The third section describes a technique to detect infected Iot devices that became a part of a botnet at the stage of compromising other devices in order to increase botnet scale. The fourth section of the article presents results obtained with the proposed method to detect botnets at the propagation stage. A conclusion contains results of implemented work.

II. BOTNETS

One of the primary intentions of cybercriminals while compromising IoT devices is to include these devices into a botnet. A botnet is a computer network consisting of infected devices controlled by malware (also called a bot) [13]. Cybercriminals employ special Trojan programs in order to

 978-1-5386-4340-2/18/$31.00 ©2018 IEEE

bypass intrusion detection systems (IDS) and intrusion prevention systems (IPS) of connected devices, gain unauthorized access and control under devices and combine them into a global network (botnet) that can be managed remotely.

Traditionally, botnets consisted mainly of compromised personal computers, but a low level of information security of IoT devices and their mass distribution led to the fact that these network devices became another attractive target for cybercriminals. Recent the most powerful attacks were performed by botnets which consisted mainly of unsecure IoT devices. The botnet Mirai is considered the largest botnet in the history, containing a huge number of compromised IoT devices [14].

The Mirai operating principle consists in the following: it performs scanning of IPv4 address space in order to find vulnerable devices with open ports TCP/23 and TCP/2323 [15], which are used by a network service TELNET [16], and then it performs a brute-force attack on these ports. In order to gain access to shell of a device Mirai employs more than 60 various combinations of default user credentials, which are disclosed publicly. As soon as a smart device becomes a part of this botnet, it starts to scan IPv4 address space in order to find other vulnerable connected devices and then compromise them. Nevertheless, infected devices continue to perform activities specified by a manufacturer, hence owners may be not aware of their IoT device being a part of the Mirai botnet and carrying out malicious actions by a command of an adversary.

Attacks implemented by the Mirai botnet laid the foundation of appearance of great number of botnets consisting of IoT devices, for example, the botnet Leet [17] and the botnet Amnesia [18]. The majority of botnets gain unauthorized access to IoT devices by performing brute-force attacks on TELNET and/or SSH [19] services. According to an examination of cybersecurity of the Internet of Things [20], nearly 400,000 IoT devices accept connections via these 2 services, moreover, some devices do not require authentication. All these IoT devices, providing they use default user credentials, can be compromised due to carried out brute-force attacks and become a part of a botnet. Nowadays a significant increase of a rate of DDoS-attacks powered by IoT devices is in evidence.

Thus, as a result of availability of a great number of weakly protected network devices, botnets are still remaining one of the major concerns in cyberspace.

In order to become an effective part of a botnet, a vulnerable network device passes through the sequence of stages shown in Fig. 1.

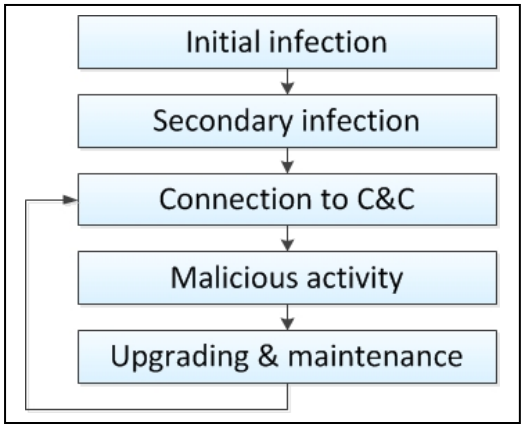

Fig. 1. A bot lifecycle.

At the first stage of the lifecycle, a compromise of a vulnerable device, considered as a potential bot, is performed. At the second stage, the malware needed to communicate with a botmaster is downloaded and installed. The third stage consists in the connection to the Command and Control (C&C) server to get instructions from the botmaster. The next stage is the stage of malicious activity, which supposes performing malicious activity on the instruction of the botmaster by the infected host. The last stage consists in upgrading and maintenance. This stage is essential for the botmaster in order to be able to effectively monitor infected hosts as long as possible, modifying their behavior by installing malware updates [21].

The break of the represented chain at any stage allows to avoid the large-scale loss caused by malicious activity performed by a botnet.

This article proposes a method to detect botnets at the propagation stage, which includes the first stage of the bot lifecycle - the primary infection.

III. METHODOLOGY

In order to detect botnets, focused on IoT devices, at the propagation stage a logistic regression model is employed in this work.

Logistic regression is a statistical model used to estimate the probability of an event based on values of a set of variables - predictors. Logistic regression is based on the logistic function $f(y)$, provided below:

$$f(y) = \frac{1}{1+e^{-y}} \tag{1}$$

In logistic regression y is expressed as a linear function of n input variables:

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n \tag{2}$$

Then, based on the input variables $x_1, x_2, ..., x_{n-1}$, the probability of an event is shown below [22]:

$$p(x_1, x_2, ..., x_{n-1}) = f(y) \quad (3)$$

In this case, the logistic regression model is used to estimate the probability that the device that initiated connection is a part of IoT botnet. In order to build the model, data about 100 botnets oriented to IoT devices and performing brute-force attacks to increase their scale was collected.

To create the logistic regression model the following parameters were selected as predictors:

- Destination port. The majority of botnets targeting IoT devices perform brute-force attacks on TELNET and/or SSH service in order to gain unauthorized access to a network device.
- Open source ports. The host sending malicious requests has at least one open port (this port is used to get instructions from C&C server).
- Number of requests. A number of requests is a number of attempts to find correct user credentials. For botnets this parameter usually has a value of 100-160.
- Even number of requests. Generally, the number of requests is even (one request contains a username and another one – a password).
- Mean interval between requests. Mean interval between requests is not very large (less than a second) with a small deviation.
- Requests on other ports. Ususally in order to gain unauthorized access to a device an infected host sends malicious packets to one open port.
- Mean size of packets. A packet size is not large enough, because it contains a default username or a password.
- Delta for packet size. The deviation from a mean packet size is not large (nearly 10 bytes).
- Mean entropy of packets. Mean entropy value lies between 2.2 and 3.8.
- Alphanumeric. Because of the fact that packets contain default user credentials, its content is usually alphanumeric.

The list of variables of the received logistic regression model, their type and weight are presented in Table 1.

TABLE I. PARAMETERS FOR LOGISTIC REGRESSION MODEL

Parameter	Type	Weight
Dst_port	boolean (22, 2222, 23 or 2323)	0.1259
Open_src_ports	boolean	0.0095
Cnt_req	integer	-0.0004
Even_cnt_req	boolean	0.2892
Interval_req	number	-0.0638
Other_ports_req	boolean	-0.5113
Packet_size	number	-0.0734
Packet_size_delta	number	-0.0093
Entropy	number	0.0107
Alphanumeric	boolean	0.0815

IV. RESULTS

According to the results obtained by employment of the proposed method to detect IoT botnets at the propagation stage, the created logistic regression model appeared to be accurate enough. To estimate the effectiveness of the proposed method, the following parameters were calculated:

- Accuracy. Accuracy (acc) is the proportion of correctly predicted events:

$$acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Where TP is true positive, TN – true negative, FP – false positive and FN – false negative.

- Precision. Precision (pre) – is the number of positive predictions divided by the total number of positive class values predicted:

$$pre = \frac{TP}{TP + FP} \quad (5)$$

- Recall. Recall is the number of positive predictions divided by the number of positive class values in the test data:

$$recall = \frac{TP}{TP + FN} \quad (6)$$

- F-Measure. F-Measure is the harmonic mean of precision and recall:

$$F - Measure = \frac{2 \cdot pre \cdot recall}{pre + recall} \quad (7)$$

The obtained results are provided in Table 2.

TABLE II. OBTAINED RESULTS FOR LOGISTIC REGRESSION MODEL

Parameter	Value
Accuracy	97.30
Precision	0.94
Recall	0.98
F-Measure	0.96

This model can be applied to detect botnets that gain unauthorized access to IoT devices by performing brute-force attacks on TELNET and/or SSH services.

V. CONCLUSION

By 2020, the number of connected devices is expected to grow exponentially to 50 billion, however, the level of information security of these devices remains low enough. That is why IoT botnets are becoming an increasingly popular instrument for performing massive DDoS attacks employed by cybercriminals.

To effectively protect against IoT botnets, a technique to detect them is essential. In this article a technique to detect IoT botnets at the propagation stage, i.e. when infected devices that are a part of a botnet compromise other devices to increase the size of the botnet, is proposed. The provided model is applicable for detection of botnets, which are propagated through brute-force attacks using the TELNET and/or SSH protocols.

ACKNOWLEDGMENT

Study of Internet of Things botnet detection techniques for information security was held within the framework of the Academic Excellence Project (Contract No. 02.a03.21.0005) of the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute.

REFERENCES

[1] Vermesan, O., Friess, P., Guillemin, P., Gusmeroli, S., Sundmaeker, H., Bassi, A., Doody, P. Internet of things strategic research roadmap. *Internet of Things-Global Technological and Societal Trends*, 2011, 1, pp. 9-52.

[2] Wu, G., Talwar, S., Johnsson, K., Himayat, N., & Johnson, K. D. M2M: From mobile to embedded internet. *IEEE Communications Magazine*, 2011, Vol. 49, pp. 36-43.

[3] Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, 2013, Vol. 29, pp. 1645-1660.

[4] Cisco, 2011, Available at: http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf (accessed 29 April 2017).

[5] Prokofiev, A. O., Smirnova, Y. S., & Silnov, D. S. Examination of cybercriminal behaviour while interacting with the RTSP-Server. 2017 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), St. Petersburg, 2017, pp. 1-4. DOI: 10.1109/ICIEAM.2017.8076437

[6] Kaspersky Lab, 2016, Available at: https://kasperskycontenthub.com/securelist-russia/files/2016/12/KASPERSKY_SECURITY_BULLETIN_2016_RUS.pdf (accessed 29 April 2017).

[7] Freier, A., Karlton, P., Kocher, P. The secure sockets layer (SSL) protocol version 3.0., 2011.

[8] Symantec, 2015, Available at: https://www.symantec.com/content/dam/symantec/docs/white-papers/insecurity-in-the-internet-of-things-en.pdf (accessed 29 April 2017).

[9] Vanhoef, M. Key Reinstallation Attacks: Breaking the WPA2 Protocol., 2017. Available at: https://papers.mathyvanhoef.com/ccs2017.pdf (accessed 29 April 2017).

[10] Gartner, 2016, Available at: https://www.gartner.com/newsroom/id/3291817 (accessed 29 April 2017).

[11] Hachem, N., Mustapha, Y. B., Granadillo, G. G., Debar, H. Botnets: lifecycle and taxonomy. 2011 Conference on Network and Information Systems Security, La Rochelle, 2011, pp. 1-8. DOI: 10.1109/SAR-SSI.2011.5931395

[12] Dobbins, R. Mirai iot botnet description and ddos attack mitigation. *Arbor Threat Intelligence*, 2016, 28.

[13] Gu, G., Perdisci, R., Zhang, J., Lee, W. BotMiner: Clustering Analysis of Network Traffic for Protocol-and Structure-Independent Botnet Detection. USENIX security symposium. 2008, Vol. 5. No. 2.

[14] Kaspersky Lab, 2017, Available at: http://www.kaspersky.ru/about/news/virus/2017/newish-mirai-spreader-poses-new-risks (accessed 29 April 2017).

[15] SANS ISC InfoSec Forums, 2016, Available at: https://isc.sans.edu/forums/diary/What+is+happening+on+2323TCP/21563/ (accessed 29 April 2017).

[16] Postel, J. Telnet protocol specification. 1980.

[17] 650Gbps DDoS Attack from the Leet Botnet. Available at: https://www.incapsula.com/blog/650gbps-ddos-attack-leet-botnet.html (accessed 29 April 2017).

[18] New IoT/Linux malware targets DVRs, forms botnet. Available at: https://researchcenter.paloaltonetworks.com/2017/04/unit42-new-iotlinux-malware-targets-dvrs-forms-botnet/ (accessed 29 April 2017).

[19] Ylonen, T., Lonvick, C. The secure shell (SSH) protocol architecture. 2006.

[20] Prokofiev, A. O., Smirnova, Y. S., Silnov, D. S. The Internet of Things cybersecurity examination. 2017 Siberian Symposium on Data Science and Engineering (SSDSE), Novosibirsk, 2017, pp. 44-48. DOI: 10.1109/SSDSE.2017.8071962

[21] Abu Rajab, M., Zarfoss, J., Monrose, F., Terzis, A. A multifaceted approach to understanding the botnet phenomenon. Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. 2006, pp. 41-52

[22] Dietrich, D., Heller, B., Yang, B. Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. 2015.