Themis: Data Driven Approach to Botnet Detection

Patrick Kalmbach, Andreas Blenk and Wolfgang Kellerer
Technical University of Munich
Munich, Germany

Stefan Schmid
University of Vienna
Vienna, Austria

***Abstract*—The detection of hosts infected with botnet malware in a timely manner is an important task, since botnets are responsible for many recent security incidents. We propose Themis, an approach based on inferring the structure of time varying IP-to-IP communication with the Stochastic Block Model (SBM). Themis use the inferred structure to detect and quantify abnormal behavior of individual hosts.**

The novelty of our approach is the use of probabilistic inference directly on host interactions to model normality. The challenges of our approach are adapting the inference process to obtain a usable output in a dynamic system, and to specify abnormal behavior with respect to the inferred structure.

Themis is able to distinguish between infected and benign hosts with accuracy larger 95 % and compares favorably against state of the art botnet detection mechanisms [1].

***Index Terms*—Cyber Security, Bot Detection, Probabilistic Inference, Unsupervised Learning, Stochastic Block Model**

I. Introduction

The Context: Detection of bots. Botnets are responsible for major network security incidents in the last two decades: Criminals take over hosts and use them to launch Distributed Denial of Service attacks, spamming or click-fraud [1]–[3]. The increasing number of hosts, especially Internet of Things devices, swell the ranks of botnets, and call for detection techniques operating in reasonable time.

The Problem: Moving Target. The design of detection techniques is challenging, as botnets keep evolving to avoid detection, invalidating previously successful methods [2], [3]. Additionally, infected hosts should be detected as fast as possible to mitigate the damage they can cause.

The Limitation: Specialization. Many bot and botnet detection approaches leverage specific assumptions. For example, specific protocols for the Command and Control (C&C) channel, or specific organizational structures [1]–[3]. Botnets can evade those approaches by changing their C&C protocol, or organizational structure. We hypothesize that a good detection approach should make as little assumptions as possible.

The Opportunity: Host communication. Our approach is motivated by the observation that hosts in a communication network can be separated into different structural groups [4], and the hypothesis that communication of infected hosts deviates from that of other hosts in their groups.

We propose Themis, a bot detection technique operating on a Traffic Dispersion Graph (TDG), in which nodes are distinctive IPs and edges represent send packages [5]. Themis captures the structure of a TDG with the Stochastic Block Model (SBM), and detects infected hosts by evaluating the likelihood of observed edges for each host in the TDG given the estimated model. Themis is a novel technique to not only detect but also quantify abnormal behavior for each host. We make only one assumption: A change in the communication pattern of infected hosts. Communication of an infected host will then have a low probability given the estimated model.

Contribution. This paper makes the case of leveraging probabilistic inference to learn a model for a TDG, and use this model to detect infected hosts. We evaluate our approach on real world traces containing both, normal and infected hosts.

Related Work. A huge body of research exists for anomaly-based intrusion detection in general [3], and botnet detection in particular [1], [2]. Closest to Themis are techniques operating on a TDG, e.g., [1] or [5].

Authors in [5] assume a specific organization to the botnet, which we do not. Authors in [1] cluster hosts based on topological node features calculated on the TDG. Hosts not in the largest cluster are viewed as potentially anomalous. Themis allows to quantify for each node how anomalous it is. Both, [1] and [5] operate on traces of multiple hours, and [1] takes hours to evaluate. In contrast, Themis is fast and uses short time windows. By inferring a model of communication in an unsupervised fashion, Themis is able to detect and quantify abnormal behavior.

Background: The Stochastic Block Model. The Stochastic Block Model (SBM) [6] is a probabilistic graphical model and represents a parametric probability distribution over graphs. The parameters are: Number of groups k, node to group assignment z, and expected number of edges θ_{rs} between nodes in group r and s. The probability of a graph $G = (\mathcal{V}, \mathcal{E})$ with multi-edges and self loops is then [6]:

$$P(G \mid \theta, z, k) = \prod_{i<j} \mathrm{Poi}_{\theta_{z_i, z_j}}(A_{i,j}) \prod_{i} \mathrm{Poi}_{\theta_{z_i, z_i}}(A_{i,i}) \quad (1)$$

where A is the adjacency matrix of G, $A_{i,i}$ gives exactly the number of self edges, and z_i gives the group membership of the i$^{\text{th}}$ node. If a graph is given and parameters are unknown, z and θ can be found by maximum likelihood: $\hat{\theta}, \hat{z} = \mathrm{argmax}_{\theta, z} P(G \mid \theta, z, k)$. Parameter k can be estimated using Minimum Description Length (MDL) [7]. The estimated parameters encode the structure of G.

Organization. Sec. II describes Themis, and Sec. III reports results obtained on dataset nine of the CTU13 corpus [8].

II. Problem and Approach

We want to detect malicious hosts during operation, i.e., we consider an online scenario. We build a series of TDGs from time windows and model the TDG of each window t as a simple un-directed graph $G^t = (\mathcal{V}^t, \mathcal{E}^t)$. Each node $v \in \mathcal{V}^t$ corresponds to a unique IP address, and each edge $(u, v) \in \mathcal{E}^t$ exists if at least one package is sent from u to v or vice versa. We do not consider sent traffic nor any other attribute.

We sequentially infer parameters of the SBM for each TDG G^t. We keep the group of previously seen nodes fixed, estimate the group of unobserved nodes, and re-estimate the expected number of edges between groups at each time step. Groups are kept fixed to prevent group switching of nodes between time steps. At each time step we calculate the log-likelihood (LL) of observed edges for each node as:

$$\log l_v^t = \sum_{r=1}^{k} \left(e_{v,r}^t \log \theta_{z_v^t,r}^t + \mid \mathcal{V}_r^t \mid \theta_{z_v^t,r}^t \right), \tag{2}$$

where $\mathcal{V}_r^t$ is the set of nodes assigned to group r, and $e_{v,r}^t$ is the number of edges from node v to nodes in group r. An anomaly score for each host is calculated as:

$$s_v^t = \mid \mathrm{med}_{z_v^t} - \log l_v^t \mid, \tag{3}$$

where $\mathrm{med}_{z_v^t}$ is the median LL from nodes in one group z_v^t. Each group provides a context against which the behavior of a host is evaluated, and groups are estimated directly from data without any prior assumptions.

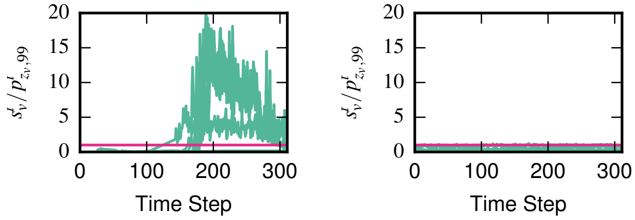
To detect anomalous hosts, we opt for a simple threshold based approach. The choice of the threshold trades detection of bots against FPR. With increasing threshold increases the risk of missing a bot, but decreases the FPR of benign hosts. We propose the 99th percentile $p_{r,99}^t$ of anomaly scores for each group r as threshold for each host in $\mathcal{V}_r^t$.

III. Evaluation

We evaluate our proposed approach on the publicly available CTU13 corpus, data set nine, since it has ten hosts that are infected with the Neris malware, and six known benign hosts [8]. We refer to the infected hosts as *bots* and the benign hosts as *normals* from here on. The bots are infected after 115 min by the authors. Before that they are benign.

We use time windows of one minute and use MDL to estimate the number of groups on the TDG of the first time window, resulting in $k = 7$, which we keep fixed. In this setting, parameter estimation for each TDG takes around 5 s. We use accuracy (ACC), false positive rate (FPR) and run length (RL) to evaluate our approach. ACC is the fraction of time windows a host is labeled correctly as benign or malicious. FPR refers to the fraction of time windows a host is falsely labeled as malicious, and RL is the number of time windows from infection to detection by Themis.

Fig. 1 plots the anomaly score of hosts divided by the respective percentile $p_{r,99}^t$. A value larger one (pink line), indicates anomalous behavior. Fig. 1a depicts the values for bots, and Fig. 1b depicts the values for normals. Fig. 1a shows a sharp increase for all bots after the authors start to infect them. Before infection, all values are well below one. For the normals values stay close to, or below one.



(a) Bots

(b) Hosts

Fig. 1: Anomaly score for bots and normals. Each score s_v^t is divided by the respective percentile $p_{v_r,99}^t$.

For bots, Themis achieves an ACC between 0.95 and 1.0, a FPR between 0.0 and 0.14 and RLs of zero and one. Themis is thus competitive to [1] using simple binary edge data.

For five out of six normals we get ACC values of 0.99 and 1.0, and a FPR between 0.0 and 0.009. For one normal results Themis in an ACC of 0.62 and a FPR of 0.37. This host is a DNS server and its behavior seems not well explainable by the estimated model. However, as Fig. 1 shows, we could decrease the FPR of normals to zero by increasing the threshold, without worsening ACC or RL for bots too much.

IV. Future Work

We plan to evaluate our proposed approach Themis on other datasets from the CTU13 corpus, and extend our focus from botnets to other malware and network anomalies.

We also intend to take past anomaly score values to detect potential bots, and want to include additional attributes on nodes and edges into the Stochastic Block Model.

An interesting avenue of research is the design of an distributed inference algorithm. This would allow Themis to scale to large communication networks, and possibly open the door for collaboration across organizational boundaries. Collaboration could increase the quality of the model while keeping privacy.

References

[1] S. Chowdhury, M. Khanzadeh, R. Akula, F. Zhang, S. Zhang, H. Medal, M. Marufuzzaman, and L. Bian, "Botnet detection using graph-based feature clustering," *Journal of Big Data*, vol. 4, no. 1, p. 14, May 2017.

[2] S. Garca, A. Zunino, and M. Campo, "Survey on Network-based Botnet Detection Methods," *Sec. and Commun. Netw.*, vol. 7, no. 5, pp. 878–903, May 2014.

[3] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 303–336, June 2014.

[4] P. Kalmbach, A. Blenk, M. Kluegel, and W. Kellerer, "Generating synthetic internet- and ip-topologies using the stochastic-block-model," in *2017 IFIP/IEEE IM*, May 2017, pp. 911–916.

[5] S. Ruehrup, P. Urbano, A. Berger, and A. D'Alconzo, "Botnet detection revisited: Theory and practice of finding malicious p2p networks via internet connection graphs," in *2013 Proceedings IEEE INFOCOM*, April 2013, pp. 3393–3398.

[6] B. Karrer and M. E. Newman, "Stochastic blockmodels and community structure in networks," *Physical Review E*, vol. 83, no. 1, p. 016107, Jan. 2011.

[7] T. P. Peixoto, "Parsimonious Module Inference in Large Networks," *Physical Review Letters*, vol. 110, no. 14, Apr. 2013.

[8] S. Garca, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods," *Computers & Security*, vol. 45, no. Supplement C, pp. 100 – 123, Sept. 2014.