International Journal of Engineering Research & Informatics (IJERI) ISSN: 2348-6481
Vol.1, Issue No 1, 2021

Dark-TRACER: A Malware Early Detection Framework Leveraging Temporal and Spatial Abnormalities

[1] P.Goutham Kumar, [2] Sala Mohan Rao, [3] Shaik Maseer, [4] A.Sai Krishna,
CSE Department,
[1,2,3,4] Assistant Professor, Dhruva Engineering Collage, Hyderabad.
Shree Engineering Collage, Hyderabad.

ABSTRACT

There is an urgent need to swiftly recognize patterns in hacking and implement appropriate defences as their prevalence rises around the world. Because there is no genuine contact taking place in the darknet, an observation and analysis of random hacks is made easier. Similar spatial patterns are seen on the darknet, where adware is spreading outbreaks through indiscriminate monitoring. Focusing on the unusual alignment of spatial patterns in darknet traffic data, we hope to solve the issue of early discovery of virus activities. Three different machine learning techniques were used in our prior research to suggest algorithms that could autonomously predict and identify unusual spatial patterns of darknet traffic in real time. In this work, we combined the previously suggested techniques into a unified system called Dark-TRACER and tested its ability to identify these malware behaviours using quantitative methods. Our large-scale darknet monitors (to /17 network sizes) were used to collect statistics on darknet activity from October 2018, through October 2020. The findings show that the techniques' flaws cancel each other out, leading to a perfect memory rate for the suggested methodology. Dark-TRACER also finds malicious activity on average 153.6 days before it is disclosed to the public by trustworthy third-party security study groups.

Keywords

Synchronization estimation anomalies, underground networks, viruses, temporal and spatial patterns.

INTRODUCTION

As more and more indiscriminate hacks have been spotted on the Internet in recent years, the expense of analysing them has risen. In order to keep the Internet safe, experts need to be able to rapidly identify global hacking patterns, identify the reasons for them, create solutions, and notify the world. This is why it's crucial to identify the actions of random surveillance attacks brought on by adware in its early stages, before an outbreak of a specific assault occurs. However, it is difficult to distinguish malware monitoring assaults from the overwhelming volume of otherwise innocuous network data. Thus, we made use of spare IP ranges. (darknets). In contrast to secret communication networks like Tor, the word "darknet" applies to surveillance networks, also known as "network telescopes." Since there is no genuine conversation taking place in the black web, the evidence of random searching is amplified. Therefore, there is a lot of useful information and not much background chatter. This makes it a useful tool for discovering patterns and themes in hacking across the world.

Dark web traffic volumes, however, have been growing at an exponential rate over the past few years. Furthermore, the goals of many interactions are unclear because we only ever see the very first messages. Multiple, unrelated cyberattacks, such as scanning activities performed for benign investigation purposes, communications with unknown causes, and misconfigured communications have all been observed concurrently in a darknet, among other things. Research should focus on differentiating these kinds of noise messages from those used in malevolent attacks. Spatial-temporal patterns of scanning are used by infected devices to corrupt new targets [1]. This is because devices attacked with identical software share scanning modules. This pattern is also seen on the deep web [2]. In this context, "spatial features" relate to the observed patterns of source addresses and target ports for messages over time. Spatiotemporal patterns are thus defined as those noticed in the temporal change of these spatial characteristics. If messages from different sites are arriving at the same target ports at roughly the same times, we say that the hosts and ports are synced.

CONCEPTUAL STRUCTURE

Dark-TRACER's general architecture is depicted in Fig. 1. Estimating the synchrony of satiate poral characteristics requires the use of three algorithms: Graphical Lasso [3], NMF [4], and NTD [5]. The components that include these algorithms are referred to as Dark-GLASSO [6], [7], Dark NMF [8], and Dark-NTD [9]. By zeroing in on seen chronicity, we can improve upon conventional malware activity detection techniques in the following ways: 1) We can lessen the impact of innocuous background communication in the darknet data and put the spotlight on the malevolent communication. 2) By detecting anomalously synchronized spatial features, malware activities that are difficult to trace by conventional manual operations can be caught before the malware infection becomes widespread, including threats that are small-scale, orchestrated, or have no visible explicit spikes. Thirdly, malware can be discovered before it expands widely if its actions are found to be in harmony with those of other malware at a time when the size of infestation is tiny. Algorithm

1 presents the Dark-TRACER framework's pseudocode. Parameters are discussed here and in Section III-C; for further information on the three methods used in this research, the reader is directed to the respective initial papers [3], [5], and [6], as well as our earlier writings [7], [8], and [9]. In light of Fig. 1 and Algorithm 1; the components are elaborated upon in the latter.

REFLECTION ON THE DATA

Dark-TRACER aims to analyse travel information from the dark web. Since routine communication is not usually noticed in the darknet (the "noise") and random scanning communication (the "signal") is watched in excess, the darknet benefits from a high signal-to-noise ratio. However, not all detected contacts in the darknet are malevolent in nature. Some of the communications noticed in the darknet are unrelated to assaults, such as legitimate monitoring activities like Shodan and Census [11],1 mysterious communication, and communications that were accidentally set up in the wrong way. Dark-TRACER is a system that can identify adware and intrusive assaults by filtering out and disregarding these types of interactions. To better comprehend international tendencies in random hacking, we have developed a comprehensive darknet surveillance system called NICTER project2. Multiple nations and institutions have deployed darknet observation systems (sensors), and roughly 300,000 IP addresses are now being tracked. There is some variation in the data collected by these darknet devices

Algorithm 1 The Framework of *Dark-TRACER*

Require: Common: T, M, t, `sensor` // *Dark-GLASSO*: γ, λ, K, θ // *Dark-NMF*: r, α, β, f // *Dark-NTD*: $\tilde{R}_n$, R_n, `epochs`, `th`
Ensure: *Alerts*
while Every t seconds in darknet `sensor` **do**
  /* Data Observation (Section II-A) */
  Darknet traffic data for T s is newly updated, then preprocess it.
  /* Spatiotemporal Feature Extraction (Section II-B) */
  Generate $V_h \in \mathbb{N}_0^{M \times N_h}$, $V_p \in \mathbb{N}_0^{M \times N_p}$, $V_{hp} \in \mathbb{N}_0^{M \times N_h \times N_p}$
  /* Algorithms and Anomaly Detection (Section II-C, II-D) */
  /** *Dark-GLASSO* **/
  if $N_h > \gamma$ **then** $V_h \leftarrow$ `random_sampling`(V_h), $N_h \leftarrow \gamma$ **end if**
  Precision matrix $\Sigma_\lambda^{-1} \in \mathbb{R}^{N_h \times N_h} \leftarrow$ `graphical_lasso`(V_h, λ)
  $d_\lambda \in \mathbb{R} \leftarrow$ `graph_density`(Σ_λ^{-1})
  $outliers_1 \leftarrow$ `anomaly_detection`(d_λ, K, θ) in *Dark-GLASSO*
  /** *Dark-NMF* **/
  $W \in \mathbb{R}^{M \times r}$, $H \in \mathbb{R}^{r \times N} \leftarrow$ `NMF`$(V_h$ or $V_p, r)$
  $outliers_2 \leftarrow$ `anomaly_detection`(W, α, β, f) in *Dark-NMF*
  /** *Dark-NTD* **/
  $outliers_3 \leftarrow$ `NULL`
  for `epochs` **do**
    $\mathcal{G}, \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)} \leftarrow$ `LRA-NTD`$(V_{hp}, \tilde{R}_n, R_n)$
    $outliers_3 \leftarrow outliers_3 |$ `anomaly_detection`$(\mathcal{G}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}, th)$ in *Dark-NTD*
  end for
  /* Issuing Alerts (Section II-E) */
  $outliers \leftarrow outliers_1 | outliers_2 | outliers_3$
  $Alerts \leftarrow$ `issuing_alerts`$(outliers)$
end while

depending on where they are and how big their telescope is. This is why Dark-TRACER performs individual analyses for each instrument. Since Dark-TRACER only considers TCP SYN packets to be attack probes, it only examines those as part of the data preparation phase. In addition, the originating server is defined as the first 16 digits of the IP address. As a result, hosts are pooled together on a more macro, institutional scale. Finally, we avoided observing commonly observed threat ports in order to emphasize the observation of previously undisclosed malware behaviours.

Quantitatively comparing detection performance for evaluation

For this purpose, we conducted an analysis of the functionality of each suggested module and described the outcomes of two trials. The first trial in this part numerically assessed the effectiveness of each module in detecting malicious actions. The second trial, discussed in Section IV, tested whether or not it was possible to identify virus behaviours in their early stages. The tcpdump program was used to capture information from the dark web, and then the R program Dark-TRACER was used to analyze the data. All tests were performed on AMD RYZEN TR 2990WX computers with 128GB of RAM at the same time, in Japan Standard Time. We produced the ground truth for a total of 35 TCP ports in this trial by directly collecting those for which malware behaviors were definitely noticed in October 2018. In order to assess the recognition accuracy, this ground truth evaluation sought to find a hyperparameter set that reduced the number of false negatives, even if some erroneous positives occurred in each module. While previous reviews of the standard technique (Change Finder) and the suggested modules (Dark-GLASSO and Dark-NMF) have been conducted and published [7], [8], this is the first evaluation of Dark-NTD using these same criteria. In the part that follows, we'll go over the dataset, how Dark-NTD was fine-tuned, and the findings of our module-by-module comparisons.

EXPLANATION OF DATASET

Both the dataset and the ground truth for assessment were taken from public ally accessible prior reports for Dark-GLASSO and Dark-NMF4. Specifically, we used data from eight darknet sensors A to H, which are spread out across the globe and have varying observational resolutions. Approximately 80,000 IP addresses make up the darknet surveillance network, with the observation size of each instrument varying from about 30,000 IP addresses (/17 subnet) to about 2,000 IP addresses (/21 subnet). Information collected in October of 2018 was used for the exercise. Sensor A, which collects the most data, sends out an

average of 81.4 M messages per day, with a data capacity of 5,605 MB. Pre-processing omitted the following 11 commonly monitored TCP ports: 22, 23, 80, 81, 445, 1433, 2323, 3389, 5555, 8080, 52869. This was done to emphasize the detection of previously undisclosed malware.

Next, in terms of specifics of the truth, Table 1 displays the TCP ports through which malware activities were analysed, as well as the traits of malware activities split up by threat category. IoT malware like Miraa, Hajime, and HNS (Hide and Seek) were the most common classification, followed by weaknesses linked to network makers and other off-the-shelf service protocols. It's well-known that the sequence number in a Miraa SYN message corresponds to the target IP address [22], [23], serving as a watermark or unique identity of the protocol. Hajime can be identified by the fact that its window size is always 14600 and that the value 0 can be found in either the high or low 1 bit of the sequence number. Many different kinds of router vulnerabilities share the fact that an HTTP link was established between the source sites that sent the scan and the router manufacturer's registration page. The same time frame saw activity on ports 5379, 6379, and 7379, as found by Cohen et al. [24]. For more information, including charts showing how often different hosts have been targeted by spyware over time, see the aforementioned prior article.

EVALUATION OF PREVENTATIVE MEASUREMENTS

Here, we evaluate how possible it is to spot malicious behaviour in its earliest stages. Specifics about the data collection,

TABLE 1. Information necessary to evaluate the viability of early discovery of malicious behaviour, such as what the situation actually is. The period of virus spread is easily visible, and the report includes 33 TCP ports and 12 kinds of threat events for malware behaviours noticed in the 17 months from June 2019 to October 2020. (RCE: Remote Code Execution, C&C: Command and Control, DDoS: Distributed Denial-of-Service, CVE: Common Vulnerabilities and Exposures, PoC: Proof of Concept).

Threat Event	TCP Port	Monitored Period and Scale at NICTER	Reveal Date	Characteristic of Observed Malware Activities
ECHOBOT [33] (Mirai Variant)	*1220,6666, 9080*	*2019-07-11 16:00, 100→200*	*2019-08-06*	A file named Richard is forced to download. More than 50 *exploits* were found, including RCE.
MOOBOT [34] (Mirai Variant)	*60001*	*2019-06-24 09:00, 50→4K*	*2019-09-27*	Moobot has a distinctive encryption method, C&C communication protocol, and infection pathway different from the original Mirai. It has the feature of sharing C&C and download servers. It actively launches zero-day attacks and DDoS attacks targeting various vulnerabilities and ports.
	9527,34567	*2019-07-11 08:00, 50→8K*		
	81,88,8000	*2019-08-30 00:00, 1K→5K*		
	82,83,85, 8081,9090	*2019-09-04 00:00, 1K→10K*		
	1588,8888	*2019-09-19 00:00, 100→16K*		
	84,5984, 8181,9200	*2019-09-21 11:00, 1K→12K*		
BlueKeep [35] (CVE-2019-0708)	*3389*	*2019-12-13 10:00, 3.5K→4.5K*	*2019-05-14*	CVE has been released by Microsoft. *window size* is fixed to *8192*.
Shenzhen TVT [36]	*4567*	*2019-12-28 22:00, 500→3K*	*2018-04-09*	A vulnerability in the Shenzhen TVT product reported in 2017 appears to have recurred. We observed the same string as then in the honeypot: D79E94C5-70F0-468D-965B-E17497CCB598. We also confirmed the Mirai feature.
MikroTik [28]	*8291,8728*	*2020-02-05 12:00, 50→1K*	*2018-08-01*	It is persistently targeted, and its *window size* is fixed at *8192*. No Mirai features were observed. We observed payloads that targeted WinBox and the API of MikroTik routers in the honeypot.
Xiongmai [37]	*9530*	*2020-02-11 12:00, 50→8K*	*2020-02-04*	A backdoor vulnerability exists in Xiongmai video recorder devices that listen on port *9530*. We confirmed the Mirai feature.
Hoaxcalls Botnet I [38] (CVE-2020-5722, CVE-2020-8515)	*8089*	*2020-03-30 07:00, 400→700*	*2020-03-24*	As for the *User-Agent* included in attack communications, XTC Botnet was frequently found; thus, it is assumed to be related to Hoaxcalls Botnet. PoC code for the vulnerabilities in Grandstream UCM6200 on 2020-03-24 and Vigor, a router manufactured by Draytek, on 2020-03-30 was disclosed.
Hoaxcalls Botnet II [39]	*9673*	*2020-04-21 04:00, 50→700*	*2020-03-09*	Similar to the above Hoaxcalls Botnet I, XTC Botnet was frequently included as a *User-Agent* in attack communications. On 2020-03-09, various vulnerabilities were disclosed in Cloud CNM SecuManager, a network management software product of Zyxel.
SaltStack [40] (CVE-2020-11651, CVE-2020-11652)	*4505,4506*	*2020-05-13 11:00, 50→300*	*2020-04-30*	On 2020-04-30, the security research organization F-Secure issued a security advisory on vulnerabilities in SaltStack Salt, an open source configuration management framework.
Linksys [41]	*55555*	*2020-06-13 15:00, 300→500*	*2020-06-13*	Communications from Iran and India have caused it and the Mirai feature to be monitored. We observed attack communications that exploited the vulnerability of the Linksys E series routers in the honeypot.
MVPower [42]	*5901*	*2020-07-13 14:00, 50→600*	*2020-07-13*	Communications from Egypt caused it and the Mirai feature to be monitored. Attack communications targeting MVPower DVR in the honeypot.
Oracle [43] (CVE-2020-14882)	*7001,7002*	*2020-10-20 00:00, 800→2.5K*	*2020-10-20*	Since 2020-10-20, we have observed a spike in the number of hosts and packets destined to *7001* and *7002*, which are the default ports of the management console of Oracle's Weblogic server.

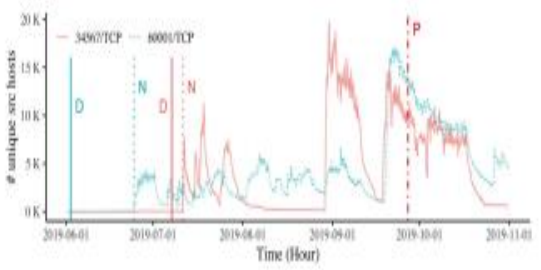

FIGURE 1. The number of unique source hosts per hour on NICTER, where Moonboot-related malware activities are observed. Here, the events that do not fluctuate synchronously with the ports in Fig. 5 are shown. D: The earliest time detected by Dark-TRACER, N: The time observed by NICTER operators, and P: The time when it was revealed to the public.

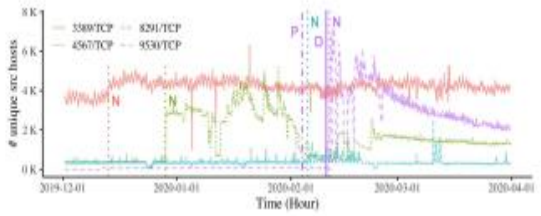

FIGURE 2. The number of unique source hosts per hour on NICTER where the malware activities were observed (December 2019 to March 2020). D: The earliest time detected by Dark-TRACER, N: The time observed by NICTER operators, and P: The time when it was revealed to the public.

experimental setup, and the assessment results are described below. This experiment included the general method of cross- validation in time series data; after learning the optimal parameters with past data in section 3, we verified them with future data in this sect

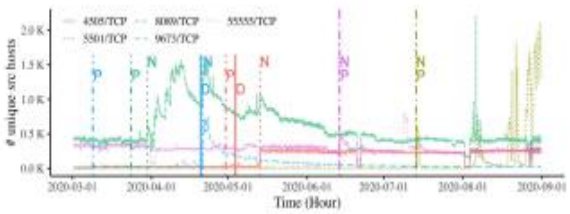

FIGURE 3. Number of unique source hosts per hour on NICTER, where malware activities were observed (March 2020 to August 2020). D: The earliest time detected by Dark-TRACER, N: The time observed by NICTER operators, and P: The time when it was revealed to the public.

DISCUSSION

In this part, we detail our framework's success and analyse our findings in depth. First, we show how Dark-TRACER can help, and then we compare all of the different modules that have been suggested in detail. The risks of our method, such as hostile assaults, and how to minimize false-positive warnings, are then discussed. At last, we provide recommendations for using Dark-TRACER in the real world. PROS OF A DARK-TRACER A. We gain the benefits stated in the preface by paying special attention to the coordination of spatial trends in darknet traffic.

1) Cleaning up garbled or out-of-sync transmissions

It's challenging to sift through darknet data and identify messages that aren't connected to attacks. Inconsistent or mysterious transmissions can cloud the results of a study of dark web activity. In this study, we zeroed in on the reality that computers hit by the same kind of software often infiltrate and monitor each other.

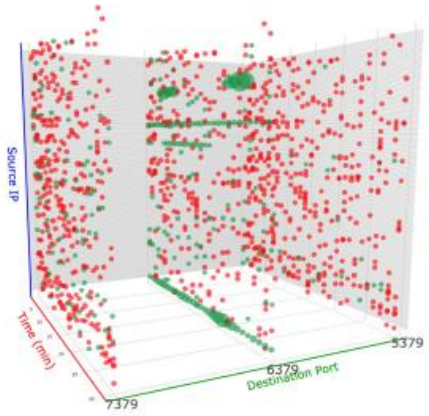

FIGURE 4. A 3D graph visualizing a case of anomalous synchronization of the spatiotemporal patterns detected from the experimental results in Section III. A scatter plot of partial Vmpp during 18:30–19:00 on October 31, 2018, is visualized. Each of the three axes is a time axis in minutes, a source-host spatial axis, and a destination-port spatial axis, and the plots represent the observed packets (element values in Vhp). For the destination-port spatial axis, there are only three points at which anomalous synchronicity was detected—5379, 6379, and 7379. Host IPs are plotted in red if they match on multiple ports within one-minute increments and in green if they do not. The red points are considered to be synchronized communications caused by factors such as malware activities, while the green points are considered to be noise communications.

in a coordinated display of space and time. It is anticipated that malevolent communications will be emphasized while noise communications will be eliminated from the analytic purview by approximating the synchrony of spatial patterns in the darknet traffic. Figure 9 provides a graphic representation of the spatial and temporal patterns' coordination for study. On October 31, 2018, between the hours of 18:30 and 9:00 PM, Dark-NMF received a warning of malware behaviour at sensor A by means of a visualization of Vhp. In this graph, the quantity of messages is displayed over time, from which server they came, and to which port they were sent. Figure 9 demonstrates that during that time frame, there were 5379, 6379, and 7379/TCP connections from the same server to those three ports. Table 1 shows that the same application, Redis, was the target of a surveillance assault on these ports. As a result, a large number of red spots will show up whenever there is an anomaly in the temporal and spatial pattern. However, the green spots can be interpreted as chaotic dialogue. By removing the green spots, Dark-TRACER is believed to be able to identify abnormalities by emphasizing the red ones (such as virus activity). (Noise communication). Between random target ports of typical darknet traffic, the red spots (synchronization between spatiotemporal characteristics) do not show as frequently as they did in Fig. 9.

TABLE 2. A comprehensive comparison of the proposed modules Dark-GLASSO, Dark-NMF, Dark-NTD.

Category		*Dark-GLASSO*	*Dark-NMF*	*Dark-NTD*
Accuracy	#FPs	**Almost None**	Many	Numerous
	#FNs	A Few	**Almost None**	**Almost None**
	Ability of Early Detection	Detection Slightly Delayed	**Detection Almost Early**	**Detection Almost Early**
	Ability of Detection for Small-scale Threats	Very Poor	Poor	**Robust**
	Ability of Detection for Constant Threats	**Robust**	Poor	**Robust**
	Ability of Detection for Short-term Threats	**Robust**	**Robust**	Poor
	Impact of Darknet Sensor Scale	Bigger is Better	**Not Dependent on Scale**	Bigger is Better
Cost	Computational Cost	High	**Low**	High
	Preprocessing	Required	**Not Required**	Required
	Alert Analysis	**Low**	High	Very High
Anomaly Detection	Pattern Decomposition	Not Performed	Performed	Performed
	Necessity of Past Data	Required	**Not Required**	**Not Required**
Spatial Feature	Host Spatial Feature	**Computable**	**Host/Port Spatial Features**	**Host/Port Spatial Features**
	Port Spatial Feature	Computationally Difficult	**can be Computed Separately**	**can be Computed Simultaneously**

RELATED WORK

We also discuss connected research on the use of darknets for measuring and detecting malicious behaviour. In addition, we provide connected research and guidance for identifying forensic monitors on darknet data, the need for which will undoubtedly arise in the near future. A. MEASUREMENT AND ANALYSIS OF THE DARKNET the darknet has garnered a lot of interest in the field of network security, and many academics are constantly studying its growth, analysis, and visualization [55]. The efficacy of darknets has been explained and the basics of different darknet setups, implementation methods, and sensor location strategies have been addressed in previous study [2, [56]-[58]. Furthermore, tracking, filtration, and categorization for the evaluation of darknets have been the subject of extensive research. In the following sections, we discuss connected research on darknet measurement analysis and Internet of Things (IoT) virus analysis. Table 8 summarizes the relevant research. Fig. As can be seen in Figure 10, the quantity of data monitored by NICTER's darknet surveillance network of 300,000 IP addresses has skyrocketed over the past few years. The 2016 appearance of the IoT virus Miraa [22] is primarily to blame. When it comes to spreading an infestation, IoT malware has a distinct advantage over traditional botnets due to its ability to monitor numerous ports simultaneously to create a large-scale botnet [48]. Furthermore, IoT malware versions exhibit competing behaviour, being frequently annihilated and reinfected over a brief time span [23]. In addition to further complicating cyber dangers, the rise of such varied and complex IoT malware makes it challenging to analyse the real status of malware tactics.

In order to swiftly and accurately identify possible dangers, it is crucial to have a method to examine IoT botnets while they are still enduring. The following research was done in different areas of darknet measurement analysis, with the exception of virus activity identification, which is covered in the following portion. By creating adware and testing strategies for eradicating fake traffic from darknets and live networks, Dainty et al. [49] added to a census-like study of IP address space usage. To probe Internet-wide scanning activities and discover patterns of widespread horizontal scanning operations, Durum eric et al. examined a large-scale darknet [50]. By pulling characteristics from massive quantities of darknet data and conducting association studies, Fatha et al. [51] developed an inference and classification tool to recognize and evaluate the exploring actions of cyber physical systems (CPS). With the help of multiple data sources, including darknet network statistics [52], Jonker et al. developed a strategy to defend against DoS assaults. One third of all /24 networks on the Internet were hit by a DoS assault in the last two years, the study revealed. By gathering IP header information from darknet traffic data, Shaikh et al. [53] were able to identify unwanted IoT devices. By comparing honeypot and darknet data, Akiyoshi et al. [54] suggested a way to identify new monitoring activities and their scope. To decipher the overall pattern of malevolent communications seen in darknets, most measurement analysis studies have been applied. Darknet data and trap-based surveillance systems like honeypotS are thus used by many studies for in-depth analysis.

CONCLUSION

To autonomously determine the synchronization of the spatial trends of darknet traffic in real time and to identify abnormalities, we proposed three separate machine learning techniques in this research. You can accomplish this with the Dark-GLASSO, Dark-NMF, or Dark-NTD techniques. We also suggested Dark-TRACER, which unifies the three approaches. Dark-TRACER was able to compensate for the shortcomings of each module, allowing it to identify all malicious behaviors while maintaining a perfect memory rate. On average, it spotted the virus 153.6 days before it was disclosed to the public by independent security researchers. We also determined that two researchers can perform the everyday tasks required to identify these risks in around 7.3 hours. The high rate of erroneous findings is currently our biggest obstacle. In this research, we found that even a basic rule-based strategy significantly cuts down on erroneous positives. Future work will focus on reducing the amount of erroneous hits by tracing the research devices' unique signatures, as detailed in Sections V D and VI-C. False negatives are a costly part of any study, so doing away with them can save money. In addition, we plan to implement an

automated version of the secondary impact analysis discussed in Section V-E to better explain the origins and specifics of the warnings picked up by Dark-TRACER. Finally, we intend to release Dark-TRACER into the wild, where it can identify dangers and malicious actions in real time and help with quick reaction.

REFERENCES

[1] G. Gu, J. Zhang, and W. Lee, ''BotSniffer: Detecting botnet command and control channels in network traffic,'' in Proc. Netw. Distrib. Syst. Secur. Symp. (NDSS), 2008, pp. 1–19.

[2] M. Bailey, E. Cooke, F. Jahanian, A. Myrick, and S. Sinha, ''Practical darknet measurement,'' in Proc. 40th Annu. Conf. Inf. Sci. Syst., Mar. 2006, pp. 1496–1501.

[3] J. Friedman, T. Hastie, and R. Tibshirani, ''Sparse inverse covariance estimation with the graphical lasso,'' Biostatistics, vol. 9, no. 3, pp. 432–441, Dec. 2007.

[4] D. Lee and H. S. Seung, ''Algorithms for non-negative matrix factorization,'' in Proc. 13th Int. Conf. Neural Inf. Process. Syst. (NIPS), 2000, pp. 535–541

[5] Y.-D. Kim and S. Choi, ''Nonnegative tucker decomposition,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2007, pp. 1–8.

[6] C. Han, J. Shimamura, T. Takahashi, D. Inoue, M. Kawakita, J. Takeuchi, and K. Nakao, ''Real-time detection of malware activities by analyzing darknet traffic using graphical lasso,'' in Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom), Aug. 2019, pp. 144–151.

[7] C. Han, J. Shimamura, T. Takahashi, D. Inoue, J. Takeuchi, and K. Nakao, ''Real-time detection of global cyberthreat based on darknet by estimating anomalous synchronization using graphical lasso,'' IEICE Trans. Inf. Syst., vol. 103, no. 10, pp. 2113–2124, Oct. 2020.

[8] C. Han, J. Takeuchi, T. Takahashi, and D. Inoue, ''Automated detection of malware activities using nonnegative matrix factorization,'' in Proc. IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom), Oct. 2021.

[9] H. Kanehara, Y. Murakami, J. Shimamura, T. Takahashi, D. Inoue, and N. Murata, ''Real-time botnet detection using nonnegative tucker decomposition,'' in Proc. 34th ACM/SIGAPP Symp. Appl. Comput., Apr. 2019, pp. 1337–1344.

[10] J. Takeuchi and K. Yamanishi, ''A unifying framework for detecting outliers and change points from time series,'' IEEE Trans. Knowl. Data Eng., vol. 18, no. 4, pp. 482–492, Apr. 2006.

[11] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, ''A search engine backed by internet-wide scanning,'' in Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur., 2015, pp. 542–553.

[12] T. Ide, A. Khandelwal, and J. Kalagnanam, ''Sparse Gaussian Markov random field mixtures for anomaly detection,'' in Proc. IEEE 16th Int. Conf. Data Mining (ICDM), Dec. 2016, pp. 955–960.

[13] A. J. Gibberd and J. D. B. Nelson, ''High dimensional changepoint detection with a dynamic graphical lasso,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2014, pp. 2684–2688.

[14] T. Idé, A. C. Lozano, N. Abe, and Y. Liu, ''Proximity-based anomaly detection using sparse structure learning,'' in Proc. SIAM Int. Conf. Data Mining, Apr. 2009, pp. 97–108.

[15] S. Liu, T. Suzuki, and M. Sugiyama, ''Support consistency of direct sparsechange learning in Markov networks,'' in Proc. 29th AAAI Conf. Artif. Intell., 2015, pp. 2785–2791.

[16] Y. Koren, R. Bell, and C. Volinsky, ''Matrix factorization techniques for recommender systems,'' IEEE Comput., vol. 42, no. 8, pp. 30–37, Aug. 2009.

[17] Q. Zhao, C. F. Caiafa, D. P. Mandic, L. Zhang, T. Ball, A. Schulze-Bonhage, and A. and Cichocki, ''Multilinear subspace regression: An orthogonal tensor decomposition approach,'' in Proc. 25th Annu. Conf. Neural Inf. Process. Syst., 2011, pp. 1269–1277.

[18] A. H. Phan and A. Cichocki, ''Tensor decompositions for feature extraction and classification of high dimensional datasets,'' IEICE Nonlinear Theory Appl., vol. 1, no. 1, pp. 37–68, 2010.

[19] A. Anandkumar, P. Jain, Y. Shi, and U. N. Niranjan, ''Tensor vs. matrix methods: Robust tensor decomposition under block sparse perturbations,'' in Proc. 19th Int. Conf. Artif. Intell. Statist., (AISTATS), vol. 51, 2016, pp. 268–276.

[20] C. F. Caiafa and A. Cichocki, ''Generalizing the column–row matrix decomposition to multi-way arrays,'' Linear Algebra its Appl., vol. 433, no. 3, pp. 557–573, Sep. 2010.