2021 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications
(GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress

2021 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics) | 978-1-6654-1762-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/ITHINGS-GREENCOM-CPSCOM-SMARTDATA-CYBERMATICS53846.2021.00034

Extracting Threat Intelligence Related IoT Botnet From Latest Dark Web Data Collection

Keisuke Furumoto*, Mitsuhiro Umizaki*, Akira Fujita*,
Takahiko Nagata†, Takeshi Takahashi* and Daisuke Inoue*
*National Institute of Information and Communications Technology, Tokyo, Japan
†Infours Inc., Tokyo, Japan
Email: furumoto@nict.go.jp, mitsuhiro.umizaki@nict.go.jp, a.fujita@nict.go.jp,
takahiko.nagata@infours.co.jp, takeshi_takahashi@nict.go.jp, dai@nict.go.jp

***Abstract*—As it is easy to ensure the confidentiality of users on the Dark Web, malware and exploit kits are sold on the market, and attack methods are discussed in forums. Some services provide IoT Botnet to perform distributed denial-of-service (DDoS as a Service: DaaS), and it is speculated that the purchase of these services is made on the Dark Web. By crawling such information and storing it in a database, threat intelligence can be obtained that cannot otherwise be obtained from information on the Surface Web. However, crawling sites on the Dark Web present technical challenges. For this paper, we implemented a crawler that can solve these challenges. We also collected information on markets and forums on the Dark Web by operating the implemented crawler. Results confirmed that the dataset collected by crawling contains threat intelligence that is useful for analyzing cyber attacks, particularly those related to IoT Botnet and DaaS. Moreover, by uncovering the relationship with security reports, we demonstrated that the use of data collected from the Dark Web can provide more extensive threat intelligence than using information collected only on the Surface Web.**

***Index Terms*—Cyber Security, Dark Web, Internet of things, Threat Intelligence**

I. Introduction

Using Tor [1], [2] to access sites on the Dark Web easily ensures the confidentiality of users on the Dark Web. Therefore, malware and exploit kits are sold on the market [3], [4], and attack methods are discussed in forums [5], [6]. Some reports indicate that Command and Control (C2) servers send commands to groups of terminals infected with malware (Botnets) set up in the Tor network [7]. Thus, on the Dark Web, some information is useful for analyzing cyber attacks, such as tools and targets that attackers use for cyber attacks. By crawling such information, information on cyber attacks that cannot be obtained using only information on the Surface Web (hereinafter referred to as *"threat intelligence"*) can be collected. However, crawling sites on the Dark Web presents technical challenges different from crawling sites on the Surface Web. For example, different image authentication methods are adopted for each site, and completely avoiding bot decision is difficult. In addition, even in a market with numerous users, sites can often be closed abruptly [8]–[10]. Furthermore, when building a crawler that collects information on the Dark Web, avoiding downloading illegal content is essential. Although datasets comprising threat intelligence on the Dark Web have been published [11], [12], the dataset in [11] contains vector data for use in machine learning for labeling illegal sites and does not include threat intelligence data. The dataset in [12] includes HTML and image files collected from the Dark Web between 2013 and 2015 and does not include data from recent years. Because trends on the Dark Web change rapidly, collecting the latest threat intelligence is crucial.

A distributed denial-of-service (DDoS) attack is a cyber attack, wherein a group of malware-infected terminals (Botnet) sends a large amount of access and data to a website or server. In DDoS attacks, the attacker sends attack commands from the Command and Control server (C2 server) to the Botnet, thereby making it difficult for the victim to identify the attacker behind the Botnet. In recent years, DDoS attacks by IoT Botnets comprising IoT devices infected with IoT malware have become particularly serious [13], [14]. IoT devices are more likely to be infected with malware because they often have vulnerable settings compared with devices having operating systems, such as Windows. With the increase of various IoT devices, such as web cameras, large-scale IoT Botnets have been constructed [15], [16]. There has been considerable research on IoT malware reporting on the types of IoT malware families and the actual status of infection [17]. Furthermore, a service exists that provides a Botnet for executing DDoS. These services are called DaaS, which stands for DDoS as a Service, also known as Booter, Stresser, and DDoSer [18], [19]. By subscribing to DaaS, DDoS attacks can be easily used even by those who do not have expert knowledge about malware and Botnets. DaaS using IoT malware Botnets has been reported [20], but no details about the means of sharing information about the service on the Dark Web have been found.

For this paper, as mentioned previously, we implemented a crawler that can solve technical challenges when crawling sites on the Dark Web. By operating the implemented crawler, we also collected information on markets and forums on the Dark Web. We focused on markets (hereinafter referred to as *"dark market"*) and forums (hereinafter referred to as *"dark forum"*) that are likely to contain threat intelligence, such as buying

978-1-6654-1762-4/21/$31.00 ©2021 IEEE
DOI 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics53846.2021.00034

and selling attack tools and exchanging information on attack campaigns. Our results confirmed that the dataset collected by crawling comprises threat intelligence that can be used to analyze cyber attacks, particularly those related to IoT Botnet and DaaS. Previous studies using only data collected from the Surface Web have primarily analyzed the damage situation and the scale of attacks by using IoT Botnet. By analyzing a large dataset obtained from the dark market and dark forum, ascertaining the actual situation of DaaS sellers and their advertising methods, buyers and their motivations, and prices and attack performance is possible. Moreover, by uncovering the relationship with security reports, we demonstrated that the use of data collected from the Dark Web can provide more extensive threat intelligence than using information collected only on the Surface Web.

II. Background

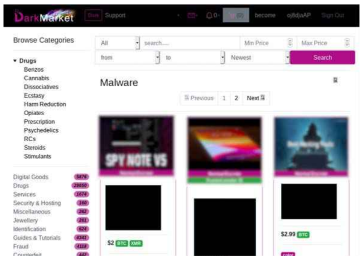
A. Dark Web

The Dark Web has various features that the Surface Web lacks. In numerous cases, a browser using Tor [1], [2], which is highly anonymous, is essential for accessing sites on the Dark Web. A wide range of sites use anonymity to buy and sell illegal goods and exchange information related to crime. Moreover, because searching the entire Dark Web is impossible, obtaining an overall picture of the sites that compose it is not easy. In this paper, we focus on markets and forums that are likely to contain threat intelligence, such as buying and selling attack tools and exchanging information on attack campaigns.

The authors of [21] proposed a machine learning method to extract threat intelligence from data collected from the Dark Web and Deep Web. Although the proposed method was evaluated using sample data, the implementation of a crawler to collect data from the Dark Web and Deep Web and the evaluation of the proposed method by using the collected data were left for future work. In [22], approximately 500 users of services related to the onion domain site were surveyed to investigate their understanding and impression of the site and their purpose of use (e.g., ensuring anonymity). In [23], authors aimed to contribute to the prevention of damage caused by intimate partner surveillance by gathering the intelligence of monitoring tools used by attackers in online underground forums on the Surface Web. In [24], using an existing dataset [12], authors proposed a method for extracting slang used in cybersecurity-related meanings in underground communication. The authors of [25] proposed a method of collecting only features extracted from images to avoid the crawler collection of illegal goods, such as child pornography.

B. Dark Market

The dark market connects multiple sellers and buyers. Numerous dark market sites require users to log in only to browse the goods, and many sites require image authentication. Various image authentication methods are used for these sites, and some sites require two-step image authentication. Many dark market sites have different popular goods. These sites prepare product categories on the site side; for example, the category of digital goods often includes subcategories of malware and exploit kits. Figure 1 shows an example of a dark market site (DarkMarket). The images of the products in the Figure 1 are blurred, and the product titles are blacked out. This page shows the list of products in DarkMarket's ”Malware” category.



[redacted]

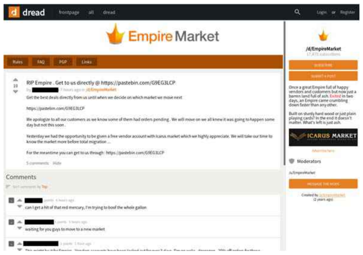
$2 BTC XMR

[redacted]

[redacted]

$2.99 BTC

Fig. 1. Examples of Dark Market Site (DarkMarket)



RIP Empire . Get to us directly @ https://pastebin.com/G9EG3LCP

by [redacted] 7 hours ago in /d/EmpireMarket

Get the best deals directly from us until when we decide on which market we move next

https://pastebin.com/G9EG3LCP

We apologize to all our customers as we know some of them had orders pending . We will move on we all knew it was going to happen some day but not this soon .

Yesterday we had the opportunity to be given a free vendor account with Icarus market which we highly appreciate. We will take our time to know the market more before total migration ...

For the meantime you can get to us through : https://pastebin.com/G9EG3LCP

5 comments Hide

Comments

[redacted] points 4 hours ago
can I get a hit of that red mercury, I'm trying to boof the whole gallon

[redacted] 1 points 5 hours ago
waiting for you guys to move to a new market

[redacted] 1 points 1 hour ago

Once a great Empire full of happy vendors and customers but now just a barren land full of ash. Exited in two days, an Empire came crumbling down faster than any other.

Built on sturdy hard wood or just plain playing cards? In the end it doesn't matter. What's left is just ash.

/u/EmpireMarket

Created by /u/EmpireMarket
(2 years ago)

Fig. 2. Examples of Dark Forum Site (Dread)

C. Dark Forum

The dark forum is a platform comprising comments from posters and other users. Many topics posted on the dark forum are related to illegal trafficking, such as drug trafficking; however, a wide variety of exchanges can be observed. Some topics pertain to cyber attacks, such as how to use malware and exploit kits. Many dark forum sites in English and other languages are available. Depending on the site, data collected from the dark forum include titles, poster IDs, ratings, comments, etc. Many dark forums do not require login with image authentication. Figure 2 shows an example of the dark forum site Dread. The IDs of thread creators and commenters in the Figure 2 are blacked out.

III. Security Threats Related IoT Botnet Focusing on the Surface and Dark Web

Figure 3 shows the Threat Model of DaaS using IoT Botnet. In the Surface Web in Figure 3, attack commands are sent from the C2 server to the IoT Botnet, which consists of IoT devices infected with IoT malware, and DDoS attacks are launched against the Target organization. It is a common practice for security vendors to observe these situations and publish them as security reports. However, it is not clear from the Surface

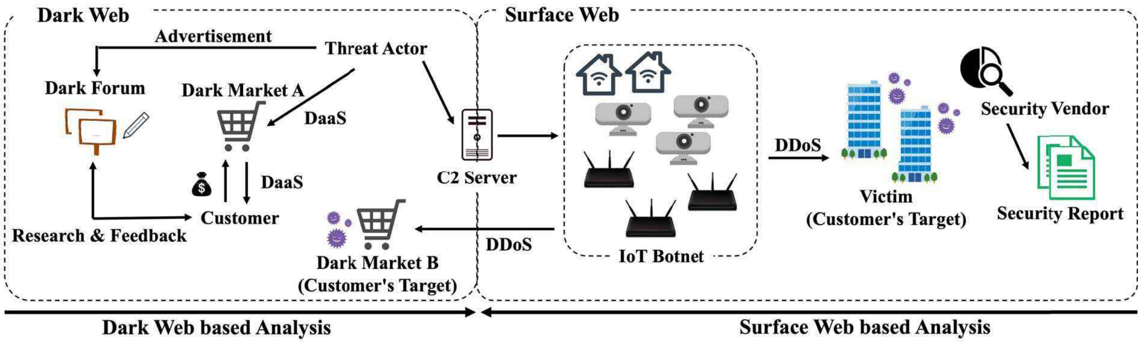


Fig. 3. Threat Model of DaaS using IoT Botnet in the Surface and Dark Web

Web based analysis whether the Threat Actor who built the IoT Botnet originated the attack or whether another Customer paid for DaaS on the Dark Web and executed the DDoS attack on the Target organization.

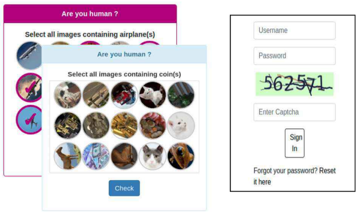
In the Dark Web in Figure 3, the Customer researches DaaS and then uses DaaS to launch DDoS attacks against the target organization. The usage scenarios of DaaS in the Dark Web are as follows:

1) Gather information about DaaS on forums.
2) Pay for a DaaS with the desired attack bandwidth and attack time on the market.
3) Execute a DDoS attack against the target.
4) Provide feedback on DaaS in dark forums and dark markets.

Figure 3 also shows the case where the target organization of the DDoS attack is a dark market. The details of this situation are discussed in detail in Chapter VI.

IV. Data Collection Strategies from Dark Web

Although there are various types of sites on the Dark Web, this paper focuses only on the dark market and dark forums, which may contain threat intelligence. However, crawling sites on the Dark Web present technical challenges different from those on the Surface Web. For example, different image authentication methods are adopted for each site, and completely avoiding bot decision is difficult. In addition, even in a market with many users, a site is often closed abruptly. Furthermore, when building a crawler that collects information on the Dark Web, ensuring that illegal downloads are avoided is crucial.


Are you human ?

Select all images containing airplane(s)

Are you human ?

Select all images containing coin(s)

Fig. 4. Examples of anti-crawling measures in the Dark Market

A. Technical Challenges in Crawling the Dark Web

Frequent Closure of Dark Web Sites. Dark markets and forums commonly experience site closures and relocations, and even popular dark markets rarely last more than 2 years. For example, EmpireMarket continues to be down, and there is a high probability that it will not be revived [26]. In this paper, we frequently update the list of URLs to be crawled. To obtain the latest onion domain, we refer to multiple sites on the Surface Web [27]–[29] that summarize the sites on the Dark Web.

Anti-crawling Measures on Dark Web Sites. Numerous sites on the Dark Web have built-in protection against crawling, such as DDoS protection and image authentication. In many cases, two types of authentications are used: one using an image and one using CAPTCHA that allows a user to enter the character string in the image. Figure 4 shows an example of the anti-crawling measures required when accessing the dark market. Because many sites require users to login simply to view products, login and image authentication support are required when crawling. For this paper, we implemented a semi-automatic crawler that assumed the minimum necessary manual login operation for sites that require image authentication.

Tor Communication Speed. The communication speed of Tor used for Dark Web crawling is very slow because Tor's communication requires triple encryption/decryption and randomly passes through onion routers installed worldwide. In this paper, we limit the collection target to threat intelligence based only on the categories for each site and respond to the extremely slow communication via Tor.

Dangers of Illegal Downloading. The Dark Web presents a risk of being charged with downloading illegal content, such as child pornography. In other words, when crawlers collect information on the Dark Web comprehensively, considering how to avoid even unintentional illegal downloads is essential. For this paper, as a countermeasure against the risk of being

TABLE I
CATEGORY AND META-INFORMATION TO BE CRAWLED IN EACH DARK MARKET SITE

Category	Security, Hacking, Botnets, Malware, Exploit Kits, Security Software, Carding, Fake Documents, Hosting, Operational Management, SOCKS, Social Engeneering, VPN, Anonymity, Cryptocurrency, Crackers, Leaks, Crypters
Meta-Information	Item ID, File Path, Page Title, Product Title, Price, Left Quentity, Sold Quentity, Category, Feedback, Description, Product Type, Meta Tags, Vendor

TABLE II
COLLECTION RESULTS FROM THE DARK MARKET

Name	Language	Lifetime[(a)]	Number of Goods	Number of Sales Vendors	Categories
ASAP Market[(b)]	English	March 2020 to now	2,884	40	7
DarkMarket	English	June 2019 to now	5,085	131	17
DarkFox	English	April 2020 to now	1,689	55	18

[(a)] Last confirmed month: January 2021
[(b)] Old name: ASEAN Market

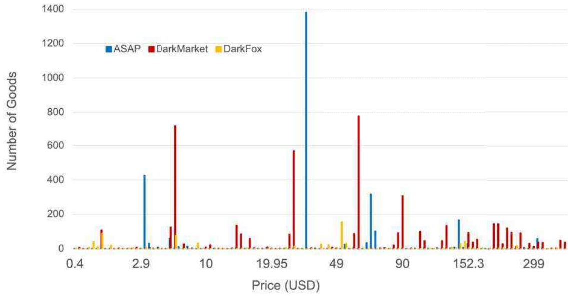


Fig. 5. Price distribution of goods in Market ASAP, DarkMarket and DarkFox

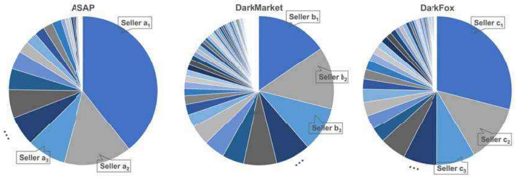


Fig. 6. Distribution of sellers according to the number of goods

charged with downloading illegal content, we collected only the meta-information of image and video files.

V. ANALYSIS OF INFORMATION COLLECTED FROM DARK MARKET AND DARK FORUM

A. Analysis of Information Collected From the Dark Market

We collected information on dark markets by operating the implemented crawlers. We selected categories related to threat intelligence for crawling (Table I). Table II shows the results collected from ASAP Market (hereinafter referred to as *"ASAP"*), DarkMarket, and DarkFox. It shows the language and lifetime of the dark market, the number of goods crawled from the dark market and the number of sales vendors and categories.

TABLE III
UNIQUE NUMBER OF IPS, DOMAINS, URLS IN THE DARK MARKET THAT MATCH THE BLOCKLIST DATABASE (NUMBER OF MATCHES/TOTAL)

Name	IP	Domain	URL
DarkFox	0 / 19	5 / 18	0 / 20
DarkMarket	2 / 81	13 / 116	0 / 148
ASAP	0 / 15	5 / 141	0 / 144

Figure 5 shows the distribution of the price range of goods in three dark market sites, ASAP, DarkMarket, and DarkFox. Figure 5 shows that many goods are priced less than $100; however, the price range in which the number of goods is listed differs between ASAP, DarkMarket, and DarkFox. Figure 6 shows the distribution of sellers according to the number of goods sold by ASAP, DarkMarket, and DarkFox. Figure 6 shows that a large percentage of the dark market is dominated by a small group of sellers. Furthermore, some ID names of sellers are used in multiple markets, thereby showing that some sellers sell many goods in multiple markets. For example, the seller with ID ”X” is ranked first in the number of goods in ASAP and seventh in the number of goods in DarkMarket. Furthermore, the user ID of the seller with the highest number of listings in market DarkFox is very similar to ”X,” and thus, it may be the same user. However, Figure 5 shows that the price range wherein the number of goods listed differs between ASAP, DarkMarket, and DarkFox; essentially, the details of the dark market listings differ depending on the site.

We extracted IPs, domains (FQDNs), and URLs from dark market descriptions and checked them against our public blocklist database (see Table IX in the Appendix A). IPs and URLs were extracted from dark market descriptions by using regular expressions, and domains were extracted from URLs. We have been building a public blocklist database by downloading IPs,

TABLE IV
CATEGORY AND META-INFORMATION TO BE CRAWLED IN DARK FORUM SITE

Category	Market, Security, Hacking, Botnets, Malware, Exploit Kits, Security Software, Carding, Fake Documents, Hosting, Operational Management, SOCKS, Social Engeneering, VPN, Privacy, Anonymity, Cryptocurrency, Crackers, Leaks, Crypters
Meta-Information	Thread ID, File Path, Category, Number of Comment, Thread Title, Post Date, Body(Comment) , Number of Vote, Author(Name, Number of Post, Number of Thread, Number of Reaction , Awards/Class, Age, Location, Website)

TABLE V
COLLECTION RESULTS FROM DARK FORUM

Name	Language	Lifetime[(a)]	Number of Threads / Comments	Number of Contributors[(b)]	Categories
Dread	English	February 2018 to now	189,260 / 1,022,857	62,203	469

[(a)] Last confirmed month: March 2021
[(b)] Total number of thread creators and commenters

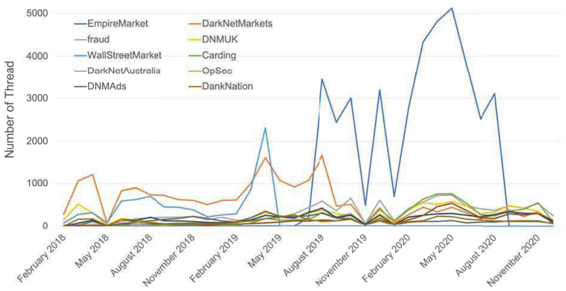


Fig. 7. Time-series transition of the number of threads by ten categories with the highest cumulative number of threads in Dread

TABLE VI
UNIQUE NUMBER OF IPS, DOMAINS, URLS IN THE DARK FORUM THAT MATCH THE BLOCKLIST DATABASE (NUMBER OF MATCHES/TOTAL)

Name	IP	Domain	URL
Dread	9 / 188	337 / 17180	0 / 4857

URLs, and domains of the public blocklist services once a day from 2021-02-26 UTC. The IPs, Domains, and URLs stored in the database as of 2021-07-15 UTC were used for the matching. Table III shows the unique number of IPs, domains, and URLs described in the dark market description and the results of matching them with the public blocklist database. Of the extracted IPs, two IPs of DarkMarket were matched with the blocklist database. Because the source of the matched IPs is a blocklist for website classification, these IPs are not necessarily malicious (e.g., IPs of websites classified as porn or violence). The product version number (e.g., 1.0.2.3) may be also extracted. As for the domain, domains matching the blocklist database were found in DarkFox, DarkMarket, and ASAP; however, the providers of the blocklists were categorizing websites or aggregating multiple blocklists. Therefore, domains are not necessarily malicious.

B. Analysis of Information Collected From Dark Forum

We collected information on dark forums by operating the implemented crawlers. We selected categories related to threat intelligence for crawling. Table IV shows an example of the selected category and meta-information collected from dark forum. The category names are replaced by different words with the same meaning. Table V shows the results collected from dark forum site Dread. It shows the language and lifetime of the dark forum, the number of comments, threads, and categories crawled from the dark forum. We crawled the posts from February 2018 to December 2020. We could not obtain information from Dread in January 2018; this is attributed to the timeframe of Dread's opening. Figure 7 shows the time-series transition of the number of threads by ten categories in Dread. These ten categories are the categories with the highest cumulative number of threads in the observation period. There tends to be a high number of threads in categories related to specific markets. This is because there are many threads where users are exchanging information about products and markets that may be closed, such as Empire Market [26].

As with the markets, we extracted IPs, domains (FQDNs), and URLs from the dark forum and checked them against our public blocklist database. IPs and URLs were extracted from dark forum posts and comments by using regular expressions, and domains were extracted from URLs. Table VI shows the unique number of IPs, domains, and URLs described in the dark forum and the results of matching them with the public blocklist database. The IPs that matched the blocklist database were provided by not only providers for website categorization purposes but also providers for use in Intrusion Detection System (IDS). As for the matched domains, similar to the domains observed in the market, some blocklists were intended for website categorization, some merged multiple blocklists, some provided a list of domains that should not be connected, and some were intended for use in the hosts file. IPs supplied by providers intended to be used in IDSs are considered to have a higher probability of maliciousness, thereby suggesting that

TABLE VII
NUMBER OF TIMES THE KEYWORD RELATED TO DAAS APPEARS IN DARK MARKETS AND DARK FORUM[(A)]

	ASAP		DarkMarket		DarkFox		Dread		
Keyword	**Count**	**Category**	**Count**	**Category**	**Count**	**Category**	**Count(Title)**	**Count(Comment)**	**Category**
Booter	7	2	12	4	3	2	7	25	6
Stresser	1	1	1	1	0	0	6	24	4
DDoSer	6	2	68	14	5	3	17	435	24
Botnet	24	5	57	12	12	5	115	672	38
DDoS	30	3	115	14	18	6	948	11256	146
C2	7	3	8	5	4	1	38	4319	167
C&C	0	0	0	0	0	0	7	51	10

[(a)] In the dark market, it is counted based on whether it is included in the product description or not. In the dark forum, the count is based on whether it is included in the thread title and comment or not. The numbers in the Category indicate how many categories the products or comments containing each keyword are distributed in.

cyber attacks in the Surface Web may be a topic of discussion in the threads containing the IPs, which indicates the possibility that a closer look could predict an attack.

VI. EXTRACTING THREAT INTELLIGENCE RELATED IOT BOTNET FROM DARK WEB DATA COLLECTION

A. Details of DaaS in Dark Markets and Dark Forums

Communications. Table VII shows that in the dark forum (Dread), there are many thread titles and comments containing keywords related to DaaS and DDoS, indicating that discussions are active. Also, threads containing keywords are distributed in many categories, suggesting that the discussion of DaaS and DDoS is distributed in various categories. Table VII also shows that few products in ASAP and DarkFox contain keywords related to DaaS and DDoS. On the other hand, in DarkMarket, there are many products that contain keywords related to DaaS and DDoS. The reason for this, which will be discussed in the next analysis on Threat Actors, is thought to be the influence of trends in each dark market. Furthermore, while Booter, Stresser, and DDoSer are used equally in the surface web reports as DaaS-related terms [18], Table VII shows that DDoSer is used more often in the dark market and dark forums.

Threat Actor. As shown in Figure 5 and 6, the distribution of the price range of products differs in each dark market, and it is likely that vendors who sell in multiple markets also change the contents they sell according to the trends in each dark market. Therefore, since Table VII shows that there are many products in DarkMarket that contain DDoSer and DDoS keywords, DarkMarket is suitable for selling DDoS tools and DaaS.

DDoS Attack Tools and Pricing. The following example shows that DaaS is often sold at a fixed attack time and bandwidth, with a different price for each performance. Table VIII shows the price range by attack bandwidth and time of the product related to DaaS. The flow of DDoS attack on DaaS in Table VIII is as follows:

1) Contact sales vendor with the IP address (website URL) of the target and the time the DDoS attack is planned.
2) At least 6 h prior to the scheduled DDoS attack time, purchase the service after agreeing on the contents through message exchange with sales vendor.
3) Performs DDoS attack against the specified target.
4) Determine whether DDoS attack is successful.

TABLE VIII
PRICE RANGE BY ATTACK BANDWIDTH AND TIME OF THE PRODUCT RELATED TO DAAS[(A)]

Price($)	Attack Bandwidth(Gbit)	Attack Time(minute)
15	15	30
20	15	60
45	60	10
55	15	240
90	60	30
150	60	60

[(a)] Examples of products by same sales vendor in categories of ”Digital Goods >Hacking” in ASAP

Advertisement. Furthermore, in Dread, the thread for product introduction has been built by the same sales vendor in the category of ASAP (Table VIII).

...We also offer DDoS attacks conducted by our botnet from 15 to 60 GB/s. We will help obtain the IP addresses of your target.

There are many threads in the dark forum for each dark market category. As shown in Figure 7, the number of threads in each dark market category is high in many cases, which indicates that users of the dark forum are highly interested in each dark market. In many cases, DaaS is also advertised in the dark forum threads in the category of the dark market where the DaaS is listed.

Customers and Victims. From the above examples of advertisements, it can be seen that users of DaaS do not need to have any expertise in malware or botnets. In addition, like the price and attack bandwidth of DaaS for the products shown in Table VIII, DaaS sold on the dark market is not attackable to large sites (e.g., e-commerce sites), but is intended for use in attacks on small-scale sites. Regarding the victims, many comments were blaming DDoS attacks by DaaS on the dark market and taking countermeasures. This is because the use of the dark market is the purpose of many users who access the dark web, and a DDoS attack that takes down the dark market would be very costly for many users.

...someone might pay them to ddos a rival market to incrase their own traffic. A ddoser might try to get money out of a market. Say pay me or ill keep up the attack.

...The people taking down markets through DDoS are usually blackmailers who will ask for high "protection payments" in order to profit themselves.

For the dark market, the downing of sites is a significant problem, and it is likely that ransom demands for DDoS attacks, like ransomware, are increasing. Therefore, although the surface web side analysis showed that the main reason for the closure of the dark market was detection by investigative agencies [8]–[10], DDoS attacks (and further ransom demands) using DaaS by peers are also considered to be a factor in the frequent closure of the dark market.

B. Collision Analysis With Security Reports on the Surface Web

Figure 7 shows many threads in the category indicating "EmpireMarket" during 2019-2020. The reason why the thread in the category indicating "EmpireMarket" decreased in 2020 in Figure 7 is that the relevant market closed [26], [30]. However, the sources on the Surface Web side do not reveal the reason for the closure of the Empire Market. As mentioned in the previous section, there are many comments in the Empire Market category thread on Dread that blame DDoS attacks on the Empire Market, suggesting that DDoS attacks are a factor in the site's closure.

Furthermore, illegally obtained credit card information was actively bought and sold on the Dark Web in 2020 [5], [6]. The number of threads in the category indicating "Carding" and "OpSec" has increased significantly since around February 2020 in Figure 7 because information exchange on "Carding" and "OpSec" was active during this period. Thus, threat intelligence that is useful for analysis in cyber attacks can be obtained by analyzing the information collected from dark markets and forums.

C. Discussion and Research Directions

In this paper, we analyzed the actual situation of DaaS by using IoT Botnet based on the data collected from the dark market and dark forum. In previous studies using only the data collected from the Surface Web, the primary analysis results were the damage situation and scale of attacks using IoT Botnet. The analysis results in this paper indicate that ascertaining the actual situation of DaaS sellers and their advertising methods, buyers and their motivations, prices and attack performance is possible. In addition, by showing the relationship with security reports, we have shown that the use of data collected from the Dark Web can provide more extensive threat intelligence than using information collected only on the Surface Web. In particular, we showed that DDoS attacks by DaaS were involved in the closure of the dark market, some of which involved DDoS attacks with ransom demands. This is a point that has not been previously pointed out in surface web side analysis. The sites on the Dark Web that collected the threat intelligence, as presented in this paper, are likely to be closed in the future, and the latest collection of major sites will be useful for future analyses and research. In this paper, we have only shown the benefit of the data collected from the Dark Web through qualitative methods. In the future, we will develop a technique for extracting threat intelligence with the help of data collected from the Dark Web by using quantitative methods based on machine learning techniques.

VII. Conclusion

Because trends on the Dark Web change rapidly, continuous collection of the latest threat intelligence is crucial. By crawling such information and storing it in a database, threat intelligence, which cannot otherwise be obtained from information on the Surface Web, can be acquired. In this paper, we implemented a crawler that can solve technical problems when crawling sites on the Dark Web. We also collected information on dark markets and dark forums by operating the implemented crawlers. Our results confirmed that the dataset collected by crawling comprises threat intelligence that can be used to analyze cyber attacks, particularly those related to IoT Botnet and DaaS. In addition, by understanding the relationship with security reports, we demonstrated that the use of data collected from the Dark Web can provide more extensive threat intelligence than the use of information collected only on the Surface Web.

Acknowledgment

This research was conducted under a contract of "MITIGATE" among "Research and Development for Expansion of Radio Wave Resources(JPJ000254)", which was supported by the Ministry of Internal Affairs and Communications, Japan.

References

[1] "Tor project: Anonymity online." [Online]. Available: https://www.torproject.org/

[2] D. McCoy, K. Bauer, D. Grunwald, T. Kohno, and D. Sicker, "Shining light in dark places: Understanding the tor network," in *International Symposium on Privacy Enhancing Technologies Symposium, vol.5134*, 2008, pp. 63–76.

[3] "Halloware Ransomware on Sale on the Dark Web for Only $40." [Online]. Available: https://www.bleepingcomputer.com/news/security/halloware-ransomware-on-sale-on-the-dark-web-for-only-40/

[4] "Scammer Uses Fake Tor Browser to Lure Victims to Supposed Dark Web Marketplace." [Online]. Available: https://www.bleepingcomputer.com/news/security/scammer-uses-fake-tor-browser-to-lure-victims-to-supposed-dark-web-marketplace/

[5] "Hackers' private chats leaked in stolen WeLeakData database." [Online]. Available: https://www.bleepingcomputer.com/news/security/hackers-private-chats-leaked-in-stolen-weleakdata-database/

[6] "Hacker Leaks 900 Enterprise VPN Server Passwords on Dark Web." [Online]. Available: https://healthitsecurity.com/news/hacker-leaks-900-enterprise-vpn-server-passwords-on-dark-web

[7] "Tor-Based Botnet Malware Targets Linux Systems, Abuses Cloud Management Tools." [Online]. Available: https://www.trendmicro.com/en_us/research/21/d/tor-based-botnet-malware-targets-linux-systems-abuses-cloud-management-tools.html

[8] "Operation Onymous — Europol." [Online]. Available: https://www.europol.europa.eu/activities-services/europol-in-action/operations/operation-onymous

[9] "Buying Drugs Online Remains Easy, 2 Years After FBI Killed Silk Road — US News." [Online]. Available: https://www.usnews.com/news/articles/2015/10/02/buying-drugs-online-remains-easy-2-years-after-fbi-killed-silk-road

[10] "Double blow to dark web marketplaces — Europol." [Online]. Available: https://www.europol.europa.eu/newsroom/news/double-blow-to-dark-web-marketplaces

[11] "DDIR: An Open Source Dataset for Darkweb Researc." [Online]. Available: https://github.com/nenaiko-dareda/DDIR

[12] "Darknet Market Archives (2013-2015)." [Online]. Available: https://www.gwern.net/DNM-archives

[13] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "Ddos in the iot: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.

[14] A. Marzano, D. Alexander, O. Fonseca, E. Fazzion, C. Hoepers, K. Steding-Jessen, M. H. P. C. Chaves, Cunha, D. Guedes, and W. Meira, "The evolution of bashlite and mirai iot botnets," in *2018 IEEE Symposium on Computers and Communications (ISCC)*, 2018, pp. 00 813–00 818.

[15] E. Cozzi, M. Graziano, Y. Fratantonio, and D. Balzarotti, "Understanding linux malware," in *2018 IEEE Symposium on Security and Privacy (SP)*. Los Alamitos, CA, USA: IEEE Computer Society, may 2018. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/SP.2018.00054

[16] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the mirai botnet," in *26th USENIX Security Symposium (USENIX Security 17)*. Vancouver, BC: USENIX Association, Aug. 2017, pp. 1093–1110. [Online]. Available: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis

[17] O. Alrawi, C. Lever, K. Valakuzhy, R. Court, K. Snow, F. Monrose, and M. Antonakakis, "The circle of life: A large-scale study of the iot malware lifecycle," in *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021, pp. 3505–3522. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/alrawi-circle

[18] "DDoS for Hire — Booter, Stresser and DDoSer — Imperva." [Online]. Available: https://www.imperva.com/learn/ddos/booters-stressers-ddosers/

[19] "The Russian DDoS One: Booters to Botnets — Botconf 2021-2022." [Online]. Available: https://www.botconf.eu/2014/the-russian-ddos-one-booters-to-botnets/

[20] "DDoS-for-Hire Service Powered by Bushido Botnet." [Online]. Available: https://www.fortinet.com/blog/threat-research/ddos-for-hire-service-powered-by-bushido-botnet-

[21] E. Nunes, A. Diab, A. Gunn, E. Marin, V. Mishra, V. Paliath, J. Robertson, J. Shakarian, A. Thart, and P. Shakarian, "Darknet and deepnet mining for proactive cybersecurity threat intelligence," in *IEEE Conference on Intelligence and Security Informatics*, 2016, pp. 7–12.

[22] P. Winter, A. Edmundson, L. M.Roberts, A. Dutkowska-Żuk, M. Chetty, and N. Feamster, "How do tor users interact with onion services?" in *In Proceedings of the 27th USENIX Security Symposium*, 2018, pp. 411–428.

[23] E. Tseng, R. Bellini, N. McDonald, M. Danos, R. Greenstadt, D. McCoy, N. Dell, and T. Ristenpart, "The tools and tactics used in intimate partner surveillance: An analysis of online infidelity forums," in *In Proceedings of the 29th USENIX Security Symposium*, 2020, pp. 1893—-1909.

[24] K. Yuan, H. Lu, X. Liao, and X. Wang, "Reading thieves' cant: Automatically identifying and understanding dark jargons from cybercrime marketplaces," in *27th USENIX Security Symposium (USENIX Security 18)*. Baltimore, MD: USENIX Association, Aug. 2018, pp. 1027–1041. [Online]. Available: https://www.usenix.org/conference/usenixsecurity18/presentation/yuan-kan

[25] E. Bursztein, T. Bright, M. DeLaune, D. M.Eliff, N. Hsu, L. Olson, J. Shehan, M. Thakur, and K. Thomas, "Rethinking the detection of child sexualabuse imagery on the internet," in *In USENIX Enigma*, 2019.

[26] "Dark web market Empire down for days from DDoS attack." [Online]. Available: https://www.bleepingcomputer.com/news/cryptocurrency/dark-web-market-empire-down-for-days-from-ddos-attack/

[27] "Deep Onion Web." [Online]. Available: https://www.deeponionweb.com/

[28] "Dark Market List 2020." [Online]. Available: https://www.thedarkweblinks.com/dark-web-links/

[29] "DeepWebSitesLinks." [Online]. Available: https://www.deepwebsiteslinks.com/darknet-markets-links/

[30] "Dark web drug haven Empire Market has mysteriously disappeared - The Verge." [Online]. Available: https://www.theverge.com/2020/8/26/21403362/empire-market-dark-web-drug-marketplace-police-shutdown-silk-road-alphabay

APPENDIX A
BLOCKLISTS

TABLE IX
BLOCKLIST PROVIDERS WE DOWNLOAD BLOCKLISTS FROM (CF. SECTION V-A, V-B).

No.	Blacklist Provider	Number of IOCs
1	squidguard.mesd.k12.or.us,	2222245
2	www.shallalist.de	1715313
3	urlhaus.abuse.ch	1439608
4	dsi.ut-capitole.fr	1237004
5	maravento/blackweb	1123895
6	lists.blocklist.de	270683
7	openphish.com	137172
8	cinsscore.com	130446
9	www.phishtank.com	58824
10	greensnow.co	28725
11	www.joewein.net	10108
12	winhelp2002.mvps.org	7313
13	myip.ms	5270
14	danger.rulez.sk	3987
15	talosintelligence.com	2215
16	sslbl.abuse.ch	663
17	iplists.firehol.org	476
18	feodotracker.abuse.ch	265
19	zonefiles.io	174
20	www.botvrij.eu	64
21	malc0de.com	21