

Development of Cyber Threat Intelligence Tool

Ahmet Yaşar Bozkus

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

SİBER TEHDİT İSTİHBARAT ARACININ GELİŞTİRİLMESİ

Ahmet Yaşar BOZKUŞ¹

¹Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Adli Bilişim Mühendisliği, Elazığ, TÜRKİYE

Özet

Teknolojinin giderek gelişmesi ve her türlü bilginin internet ortamına taşınması yeni sorunlara sebep olmuştur. Sorunların başında bu verilerin güvenli bir şekilde saklanması ve muhafaza edilmesi gelmektedir. Çalışmada siber güvenlik ürünlerinden bahsedilmiş ve ne işe yaradıkları belirtilmişlerdir. Siber saldırılarının boyutları anlaşılması için saldırıya uğrayan büyük şirketlere yer verilerek yapılan büyük saldırılara değinilmiştir. Siber güvenlik ürünleri tek başına yetmemekte ve siber güvenlik operasyon merkezinde çalışan analistlerin siber tehdit istihbarat ihtiyacı olmaktadır. Siber tehdit istihbaratı için kullanılabilecek ücretli veya açık kaynak kodlu çözümler bulunmaktadır. Fakat istihbarat kaynaklarının çıktıları birbirini tutmamaktadır. Global ve yerli çözüm istihbarat kaynaklarından bahsederek bunlar içerisinden ücretsiz şekilde paylaşım yapan kaynaklardan veri kazıma yöntemi ile ip, hash ve domain bilgisi sorgulayan bir araç geliştirilecektir. Bu sayede tek yerden birden fazla siber tehdit istihbaratı aracının çıktısının görmesi sağlanacaktır. Yapılan sorguların sayısının fazla olması durumunda ise bir excel dosyasından okuma yaparak okuduğu değerleri tek tek sorgulayıp farklı bir excel dosyasına kaydedebilecek şekilde bir özellik eklenmiştir. Mevcutta bulunan siber tehdit istihbaratı araçlarında olmayan bir özellik olarak hash sorgusu yapıldığında kullanılan araçlardan dosya ismi bulunduğu taktirde bu dosya exe veya dll uzantılı ise farklı sitelerde ne işe yaradığı konusunda kazıma yaparak bir bilgi bulduğunda program linkini paylaşmaktadır.

Anahtar Kelimeler: Siber Güvenlik, Siber Tehdit İstihbaratı, Açık Kaynak ile Siber Tehdit İstihbaratı

Manuscript Title

Abstract

The gradual development of technology and the transfer of all kinds of information to the internet caused new problems. The most important problem is the safe storage and preservation of this data. In the study, cyber security products were mentioned and what they did was specified. In order to understand the extent of cyber attacks, large attacks on large companies that have been attacked are mentioned. Cyber security products alone are not enough and analysts working in the cyber security operations center need cyber threat intelligence. There are paid or open source solutions that can be used for cyber threat intelligence. But the outputs of intelligence sources do not match. By talking about global and domestic solution intelligence sources, a tool will be developed to query ip, hash and domain information with the method of data scraping from the sources that share them free of charge. In this way, it will be ensured that the outputs of more than one cyber threat intelligence tool can be seen from one place. If the number of queries made is high, a feature has been added so that it can read from an excel file and query the values one by one and save them in a different excel file. As a feature that is not available in the existing cyber threat intelligence tools, when a hash query is made, if the file name is found from the tools used, if this file has an exe or dll extension, it shares the program link when it finds information about what it does by scraping on different sites.

Keywords: Cyber Security, Cyber Threat Intelligence, Cyber Threat Intelligence with Open Source

1 Introduction

Today, technology is advancing at a very high speed, and it is capturing our lives in a very high way. Individuals want to have the latest technological smart devices and share their lives in social media applications. This situation causes a violation of people's life and privacy. The companies that own these applications are trying to ensure the safety of their customers and companies without stopping in the background. While wars were fought with swords and arrows in ancient times, the discovery of gunpowder brought the war environment to a different dimension. Today, the point where technology has evolved has brought wars to the cyber realms. With cyber wars, the systems of countries can be rendered ineffective and the systems of institutions can be rendered dysfunctional. As in the past, the most important sources have been the sources of intelligence today. By using, collecting and processing these resources well, the attacks to be made or the attacks that have been made can be detected. In our study, it is aimed to make an intelligence query from a single source using existing sources. In addition, an important issue in cyber intelligence is domestic intelligence. By not using domestic intelligence, it may cause important information to go to foreign companies or go to the states.

2 Cyber security

It is a concept that has been shaped with the development of technology and cyberspace, which is a part of our lives. In a study conducted by Kramer, it was seen that 28 different definitions were made for cyber security [1]. Cybersecurity is the cornerstone responsible for the security of a connected world over the Internet. The unprecedented expected increase in the number of Internet users, data and devices around the world in the coming years will bring with it great opportunities but equally daunting challenges.

2.1 Cyber security solutions

It cannot be finished by counting the damages that can be caused in the field of cyber security. As our digital life and the operation of our businesses on digital platforms gained momentum, cyber security started to take a very important place in our lives. According to recent studies in the field of cyber security, many critical data such as company information and offers, remote monitoring, medical information, financial data are stolen and used in

this field [2]. Increasing attacks provide ample reason to protect financial data, personal data, corporate and government data, health data, and any data that matters. Cyber security solutions aim to prevent such situations. Thanks to these solutions, we can protect the system with the least damage by intervening in an informed manner during or after the attack. The important ones of cyber security solutions are given below.

2.1.1 Antivirus

Antivirus programs have been trying to protect computers against viruses since they were first developed. With the development of communication networks in recent years, viruses can spread quickly and easily, infect many computers and harm users [3]. Antivirus programs can work with hash matching. The hash information of the pests previously analyzed by the analysts is added to the antivirus programs and the antivirus program deletes it when it detects it. New generation antiviruses are being developed in a way that can perform behavioral analysis like EDRs.

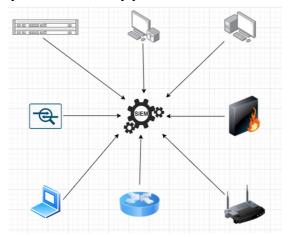
2.1.2 Endpoind detected and response

EDR tools detect known intruder behaviors and detect intruders or attempts to do so [4]. EDR tools are advanced versions of antivirus programs. Traditional attack methods are no longer used. Now, special attacks are carried out on institutions, individuals, companies and states. Since these attacks are carried out specifically, anti-virus programs are insufficient [5]. In such scenarios, EDR solutions come into play. By performing behavioral analysis, EDR can detect attacks that attackers have made or can do in line with the rules written before. The attacks carried out via powershell.exe and cmd.exe, which are considered legal by Microsoft and used for that purpose, are detected by EDRs [6].

2.1.3 Security information and event menagement

It is one of the products that has become increasingly popular and important in our country lately. The management of cyber security operation centers is one of the products they make. The SIEM product can directly collect logs on existing security devices (firewall, ips, ids, waf, etc.), logs of network devices (router, access point, switch), server logs and client logs as shown in figure 1. The structures of SIEM products may vary according to the manufacturer. While some companies collect the

logs on a different device and then send them to the SIEM product, some companies can send them directly to the SIEM product. Rules and correlations are written on SIEM products. In line with the written rules and correlations, suspicious activities are seen as alarms on the interfaces of SIEM products. The resulting alarms are analyzed by analysts and necessary precautions can be taken.



Şekil 1. SIEM log collection

2.1.4 Security orchestration automation and response

SOAR products aim to ease the burden of cyber security operations center employees and facilitate management [7]. SOAR enables products to be managed from a single place by integrating security solutions. SOAR can collect and display alarms from multiple SIEM products. It can make changes by interfering with security devices remotely and managing them. By writing rules that allow it to do this automatically, in case of an attack you specify, it can instantly take an axiom on the relevant security device.

2.2 Dimensions of cyber attack

It was seen on September 11, 2001 how much damage cyber attacks can cause. The cold war had now left its place to a new type of war. By removing the borders of the countries, the military time and place rules would be disabled [8].

2.2.1 WannaCry

WannaCry, also called ransomware, is a large-scale cyber attack that infected 230,000 computers in 99 countries around the world in May 2017, demanding ransom in 28 languages in 28 languages. This malware targets Microsoft Windows operating systems. Microsoft sent an update 2 months before the attack started, but users or device administrators were exposed to the attack because

they did not update regularly [9]. WannaCry encrypts files on the infected computer, making it inaccessible. If the ransom was not paid within the specified time, the requested ransom amount would be increased. However, if the payment was not made, it would delete the files on the computer. Passwords were cracked with mistakes made on WannaCry. Later new versions have minimized the chance of data recovery by using advanced encryption and propagation algorithms [10].

2.2.2 Stuxnet

It was thought that places with closed-circuit internet were very safe and would not be infected with malware since there was no external connection. The Stuxnet virus is thought to have been transmitted to Iranian nuclear facilities via a USB memory stick. The virus, which was discovered in June 2010, was defined by cybersecurity experts as "an advanced computer program designed to penetrate and control remote systems in a semi-autonomous manner" [11]. Stuxnet is a worm virus that infects Siemens industrial software running on Microsoft Windows operating systems [12]. Worm viruses can spread to unprotected computers very quickly by copying themselves over the network or internet connection in the system they enter [13].

2.2.3 Mirai

In today's technology, every device is getting smarter. Smart vacuums, cars, phones, washing machines. watches. wristbands, televisions, etc. Devices like these are getting smarter day by day. These devices are aimed to make our lives easier by making them smart, but here the question of how the devices are against cyber attacks as much as they are smart comes to mind. Botnet networks used to take over computers connected to the Internet and turn them into zombie computers. Here, smart devices and IoT devices that entered our lives attracted the attention of attackers by connecting to the internet and not containing antivirus programs.

2.3 Large companies hacked

Cyber attacks are increasing day by day. Companies employ expert teams to counter them. These teams monitor systems 24/7 and use products produced for cyber attacks. Even in such cases, companies that attach great importance to this business can fall into the targets of attackers. Attackers attack companies and institutions by constantly developing new attack methods.

2.3.1 Linkedin

LinkedIn, which is one of the most popular social media applications today, has been exposed to cyber attacks in the past years. The attack was carried out on June 5, 2012 and the account information of approximately 6.5 million users was stolen [14].

2.3.2 JP morgen

In September 2014, they announced that they had been attacked. The attack was noticed in July 2014 and stopped in August. Although the user's login information, passwords, and social security numbers were not captured in the attack, the account holders' names, phone numbers, and e-mail addresses were captured.

2.3.3 Sony

A hacker group called "Guardians of Peace" managed to infiltrate the systems of the movie studio Sony Pictures on November 24, 2014. In the attack, the attackers seized information such as personal information of employees and families, copies of Sony films that have not yet been released, and plans for future films. They then infected the malware named Shamoon wiper to wipe Sony's database.

2.3.4 FireEye

It is a US-based cyber security firm. FireEye provides cyber security services to many countries including Turkey. FireEye said in a statement that it was attacked by a state-sponsored APT group. He reported that cyber attack tools belonging to FireEye were found in the attack [15].

3 Cyber threat intelligence

Cyber threat intelligence is a type of intelligence that helps to detect the "motivations", "purposes" and "methods" of attackers by collecting data by following different sources of threats that may occur or that may occur against institutions and organizations, processing them with certain algorithms. Cyber threat intelligence is a field of cyber security that focuses on finding, collecting and analyzing data about current and potential attacks that threaten the security of institutions, companies and individuals. The benefit of cyber threat intelligence is that it saves financial costs by preventing data breaches that may occur. Its purpose is to provide security by showing the threats that may occur to institutions and

organizations, allowing real-time axioms to be taken [16].

3.1 The importance of cyber threat intelligence

Cyber threat intelligence is used to protect an organization or its existence. It is important to detect possible attacks and vulnerabilities early. Cyber threat intelligence is as important as intelligence in wars in ancient times. You can take advantage of the attacks by providing the necessary intelligence information, or you can be informed and prevent future attacks with the necessary intelligence. It is essential for ensuring cyber security.

3.2 Cyber intelligence resources

Cyber threat intelligence is generally provided by foreign companies. These companies make reports of newly found vulnerabilities by sharing attacker IP addresses, malicious domains and programs. These tools can be paid, free or some of them free. Cyber threat intelligence is also provided by open source software. As we aim in this project, it is possible to write your own intelligence tool by combining these tools and collecting free tools, information on github and darkweb and making it meaningful.

3.2.1 Virustotal

VirusTotal is a widely used intelligence tool by cybersecurity analysts. VirusTotal was developed as internet and browser-based security software in 2004 and later acquired by Google [17]. Named among VirusTotal 2007's top 100 products by PC World magazine. Developed by Hispasec Sistemas, VirusTotal is a platform that has an independently IT security laboratory and uses many command line versions of antivirus programs and is regularly updated with official updates released. Currently, there are 74 antivirus programs on VirusTotal [18]. IP address, hash information, file name and file can be uploaded via VirusTotal and scanned to this antivirus program by users and their vulnerabilities can be questioned. If you query an IP address via VirusTotal, you can see the malicious programs associated with that IP address. At the same time, it is possible to access the comments made by the users.

3.2.2 IBM X-Force

IBM (International Business Machines) is the world's largest information technology company, headquartered in Armonk, New York, USA. Operating in more than 170 countries, IBM has

more than 410,000 employees. IBM, the company that receives more new patents every year around the world, operates in many areas. These; computer and hardware production, services, software, server services and R&D [19].

IBM also stands out with its solutions in the field of cyber security. IBM X-Force Exchange is a cloud-based threat intelligence platform that enables you to use, share and act on threat intelligence. It allows you to quickly research the latest global security threats, gather actionable intelligence, consult experts, and collaborate with other users. Powered by human and machine-generated intelligence, IBM X-Force Exchange leverages IBM X-Force scale to ensure users are aware of emerging threats [20].

IBM X-force offers web threat monitoring of over 25 billion web pages and is backed by a database of over 96,000 vulnerabilities. You can query IP address, URL, application, hash and vulnerability. The inquiries you make are scored out of 10 and categorized. As a result of the query, we can see in detail information such as in which year it was marked as risky and when it was removed.

3.2.3 AbuseIPDB

AbuseIPDB is a platform where we can query url and IP address. Here, risky IP addresses are shared with their attacks. When an IP address search is made, it is seen that which user and which country is reported as malicious. The threat score on the site is given out of 100. This score varies with concepts such as the number of notifications made and the nature of the attack in the notifications. It can be blacklisted by querying from this site and collecting intelligence about the incoming IP [21].

3.2.4 Cyber intelligence in Turkey

The importance of cyber threat intelligence is increasing day by day in our country. People working in the field of cyber security make heavy use of the above-mentioned or unspecified cyber threat intelligence sources. However, these products are generally of foreign origin. This is very important for a country because when you use a foreign intelligence tool, you have to provide information such as what you are looking for on that site or what is important to you. In fact, when you want to provide cyber intelligence here, you are giving intelligence weakness. Studies on cyber threat intelligence are carried out in our country. Although the developed cyber threat intelligence tools have reached a certain level of maturity, they are also used.

3.2.5 Domestic cyber intelligence U.S.T.A.

It is a cyber threat intelligence tool with a web-based management screen developed by the national cyber threat network invuctus company. It is an R&D project approved by Technopark Istanbul and started to be developed in 2012, a subsidiary of the Undersecretariat of Defense Industry. Within the scope of this project, the techniques, tools and procedures in which cybercriminals operate are determined and the intelligence obtained is sent to the relevant institution via the platform. As Turkey's first and only world's leading cyber intelligence platform, USTA has achieved many successes internationally and was awarded the "Best Security Product of the Year" award in 2015[22].

3.2.6 Indigenous cyber intelligence nebula

Cyber intelligence has become the most important in information security issues. Manufacturers' product-based intelligence services often fail to provide adequate security, especially in local attacks. The Nebula Cyber Intelligence Tool provides intelligence information under three different headings as Threat Intelligence, System Intelligence and Brand Intelligence, and its webbased interface provides management and reporting from the management screen[23].

4 Material and method

This work has one main purpose. This purpose is to obtain risk score results on the platform we have written by using python, by collecting the sites that provide free cyber threat intelligence and the lists on github, by performing data scraping with the selenium and requst library. This process will take place in two stages.

4.1 Data collection phase

Python is an easy-to-learn programming language because it is far from machine language, close to human language, and it supports many standards and provides easy integration with other popular programming languages such as C and C++. It is particularly in demand among beginners due to its ease of use and the availability of many resources. According to the research of Northeastern University, python is the most popular programming language in 2020 [24]. We will use 2 different libraries to scrape data on Python. With the Request module, one of them, requests on the web can be easily managed. With this module, you can send HTTP requests such as GET, POST, PUT

and DELETE to API endpoints. You can receive the responses to the sent requests and extract the data you want. However, some websites are able to prevent this. For such cases, the selenium module is used. Selenium works in multiple programming languages and browsers. Selenium allows you to open the browser and go to the site you want and create bots to test.

4.2 Data parse stage

Collected data comes to us as html. When normal users want to view these codes on the internet browser, it is possible to display them as shown in Figure 2 by pressing the F12 shortcut key from the keyboard.



Şekil 2. Page source

Thanks to the modules it contains, the selenium library makes it easy to reach the desired data. Since the Request library does not have such a feature, help is obtained from the library called beautifulsoup4. After the beautifulsoup4 library makes the data downloaded by the Request library meaningful, the desired data can be retrieved with the help of the "find" and "find_all" commands. The risk scores extracted from the captured data are printed on the screen with the "printf" command as seen in Figure 3.



Şekil 3. Uygulama ekran çıktısı

5 Results

It has been observed that the cyber threat intelligence tools used are different from each other. These differences are shown in Table 1.

Tablo 1. Comparison of cyber threat tools

IBM	Virus Total	AbuseIPDB
Xforce		

		T	I
Hash Risk Inquiry	Yes	Yes	No
IP Risk Inquiry	Yes	Yes	Yes
Domain Risk Inquiry	Yes	Yes	Yes
URL Risk Inquiry	Yes	Yes	Yes
File Risk Scan	No	Yes	No
Risk Score	Yes	Yes	Yes
Seeing Who and From Which Country the Notifications Are Made (Without Login)	No	No	Yes
Seeing the Years of Notifications	Yes	No	Yes
Risk Score Display Range	0-10	Shows Change.	0-100

As can be seen in the table, these platforms have advantages and disadvantages from each other. By developing the cyber intelligence platform, which is the aim of the seminar, as can be seen in the image shown in Figure 3, the same IP address is seen as having different risk scores on different platforms. Using a single platform here can lead to unreliable results.

6 Conclusion

In this seminar, a new threat intelligence tool was developed with python programming language by collecting data from existing cyber threat intelligence tools and open source github. The following conclusions have been reached regarding cyber security and cyber threat intelligence in our country and in the world.

Cyber security companies are developing new products against the increasing cyber attacks today. It is important to constantly update these developed products. If the updates are not made, it may be

exposed to the attack of the large companies whose examples are given in the seminar.

If the issue of cyber security is not given importance, it can have great financial losses. As mentioned in our seminar, related companies can be held responsible for stealing information such as credit cards of users of companies.

Although companies that provide cyber security services or provide cyber threat intelligence play a major role in cyber defense, hacking them leads to the conclusion that it is not a completely secure system.

Intelligence has a very important place in wars from history to the present. Today, this situation continues and a new intelligence called cyber threat intelligence has emerged. It is important to reduce the risks posed or to be caused by the attacks. However, in general, it has been seen that intelligence sources are of foreign origin and it has been seen that companies, institutions and individuals provide intelligence to the relevant intelligence tools while they want to receive intelligence data.

Finally, it was observed that intelligence tools gave different results in the interrogations made as in Figure 3. While one intelligence tool specifies the same IP address as risk-free, another intelligence tool specifies the same IP address as risky. Here, it is seen that a decision should be made to collect data from the source as much as possible and examine all data. At the thesis stage, the application developed in the seminar will be developed and made to work on the web interface. In addition, query diversity will be increased by adding domain, url, hash, file query information. With Selenium, it will be brought into a structure that can be searched with certain keywords by the intelligence tool with a web interface that collects data on the dark web and saves the data it collects. If the users enter the keys they want in the logins made through the membership system, the bots collecting data on the dark web will find data matching these words, and the user will be sent the address of the relevant site.

Thank

I would like to thank my valuable advisor Assoc. Dr. I would like to express my endless thanks to Fatih ERTAM.

I would like to thank the Fırat University Scientific Research Projects Coordination Unit (FÜBAP) for their support of this thesis.

Resources

- [1] F.D. Kramer, S.H. Starr, L.K. Wentz, Cyberpower and national security, 2011.
- [2] https://www.deskport.com.tr/cozumlerimiz/siber-guvenlik-cozumleri (01.06.2021)
- [3] K. Wamura, (12) United States Patent, 2 (2012).
- [4] W.U. Hassan, A. Bates, D. Marino, Tactical provenance analysis for endpoint detection and response systems, in: Proc. IEEE Symp. Secur. Priv., 2020: pp. 1172–1189.
- [5] https://www.cozumpark.com/endpoint-guvenlik-cozumlerinin-geleceginde-edr-ve-mdr (01.06.2021)
- [6] R.D. Han, C. Yang, J.F. Ma, S. Ma, Y.B. Wang, F. Li, IMShell-Dec: Pay More Attention to External Links in PowerShell, in: IFIP Adv. Inf. Commun. Technol., 2020.
- [7] C. Islam, M.A. Babar, S. Nepal, A multi-vocal review of security orchestration, ACM Comput. Surv. (2019).
- [8] https://www.nato.int/docu/review/tr/articles/20 11/09/04/yeni-tehditler-siber-boyut/index.html (01.06.2021)
- [9] https://www.kaspersky.com.tr/resourcecenter/threats/ransomware-wannacry (01.06.2021)
- [10] M. Akbanov, V.G. Vassilakis, M.D. Logothetis, Ransomware detection and mitigation using software-defined networking: The case of WannaCry, Comput. Electr. Eng. (2019).
- [11] C. Stevens, Assembling cybersecurity: The politics and materiality of technical malware reports and the case of Stuxnet, Contemp. Secur. Policy. (2020).
- [12] S. BIÇAKCI, NATO'nun Gelişen Tehdit Algısı: 21. Yüzyılda Siber Güvenlik, Uluslararası İlişkiler Derg. (2014).
- [13] https://www.kaspersky.com.tr/resource-center/threats/viruses-worms (01.06.2021)
- [14] P. Doucek, L. Pavlíček, J. Sedláček, L. Nedomová, Adaptation of password strength estimators to a non-english environment—the Czech experience, Comput. Secur. (2020)
- [15] https://www.cozumpark.com/fireeye-hacklendi (01.06.2021)
- [16] https://www.bgasecurity.com/2019/05/sibertehdit-istihbarati-nedir-bolum-1 (01.06.2021)
- [17] https://www.webtekno.com/virustotal-android-apk-indir-h87680.html (01.06.2021)
- [18] https://support.virustotal.com/hc/enus/articles/115002146809-Contributors (01.06.2021)
- [19] https://tr.wikipedia.org/wiki/IBM#:~:text=IBM% 20(International%20Business%20Machines%3B% 20Uluslararas%C4%B1,en%20b%C3%BCy%C3%B Ck%20bili%C5%9Fim%20teknolojisi%20%C5%9Firketidir.&text=Kart%20delici%20makinelerin%20 bulucusu%20Herman,Watson%20taraf%C4%B1nd an% (01.06.2021)

- [20] https://www.ibm.com/tr-tr/marketplace/ibm-xforce
 - exchange#:~:text=Tehdit%20istihbarat%C4%B1%20%C3%BCzerinde%20ara%C5%9Ft%C4%B1rma%20ve,tabanl%C4%B1%20bir%20tehdit%20istihbarat%C4%B1%20platformudur (01.06.2021)
- [21] https://some.saglik.gov.tr/Files/Dokumanlar/Fark indalikDokumanlari/SOME_Kurumsal_Aglarda_Alin acak_Onlemler.pdf (01.06.2021)
- [22] https://www.siberistihbarat.com (01.06.2021)
- [23] https://www.nebulabilisim.com.tr/urunler/nebula -siber-istihbarat-servisi (01.06.2021)
- [24] https://www.northeastern.edu/graduate/blog/mo st-popular-programming-languages (01.06.2021)