Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

Bruteforce In The Hydra Process And Telnet Service Using The Naïve Bayes Method

Tarisno Amijoyo[1], Rusydi Umar[2], Anton Yudhana[3]

[1,2,3]Magister Teknik Informatika,
Universitas Ahmad Dahlan, Jl. Kapas No.9, Semaki, Kec. Umbulharjo, Kota Yogyakarta, Daerah Istimewa Yogyakarta 55166, Indonesia

Email: ahbibadil@gmail.com[1], rusydi_umar@rocketmail.com[2], eyudhana@mti.uad.ac.id [3]

ARTICLE INFO

Article history:
Received: 12/04/2020
Revised: 23/04/2020
Accepted: 01/05/2020

Keywords:
BruteForce,
Hydra,
Telnet,
Naïve Bayes

ABSTRACT

Internet technology is increasingly developing, in accordance with the needs of an increasingly millennial society. The growth of internet technology has been accompanied by an increasing variety of attack techniques aimed at the media. BruteForce is one of the techniques of attack (hacking) against the internet network by hacking passwords from users who are active or inactive at that time. Through a process with a small application such as Hydra accompanied by telnet service, then hacking techniques with BruteForce can be run, to attack an internet network. The Hydra process and the telnet service referred to in the study are those that included a "wordlist" file containing user information and password data. The count to detect, whether entering the network is a BruteForce attack or just normal / normal network traffic, can be done with the Naïve Bayes theorem or the Naïve Bayes method. The purpose of the Naïve Bayes method, is to be able to distinguish and classify, which is attack traffic, which is normal / normal network traffic. The reason for using the Naïve Bayes method is that the formulation in this method is easier to calculate when classifying the type of network traffic. From the results of the classification, it can be used as a reference for a network administrator to be more careful and better understand how to protect and prevent attacks from hackers, especially those who use BruteForce hacking techniques. The hope of this research, can be understood and understood so that it can be very easy for a network administrator to distinguish between the two traffic on the network, whether it is attack traffic and / or normal / normal traffic that is through the display on the Wireshark tool.

1. Introduction

The virtual world or the internet is very advanced and continues to grow. The ease of accessing or browsing the internet, as it is today, is what allows continuous progress on the internet media. So it is not impossible if some domain / website addresses often become the media to be attacked / hacked. As said by the Information and Communication Technology Infrastructure Development Division, Djoko Purwono, Thursday (06/07/2012), that most websites in Indonesia are hacked by hackers consisting of 50% of government-owned websites and 50% of private agencies which all of them indicated because the user is fad, the user feels dissatisfied with the website, or even the user wants to be famous and known to others. Recently several accounts on social / social websites have been hacked by hackers and then sell accounts that have been hacked. Bukalapak is one of the Indonesian websites that was victimized by the hacking [1]. Website is generally used to store data, information needed by the owner and by other information seekers, because it is very reasonable that the internet media is often a destination for the general public in seeking information and including for people who want to do evil. Often we hear and read about a variety of many criminal acts against other people, other companies, against government agencies via the internet, which is often said by the term cybercrime [2].

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

Techniques to attack a website address, with the intention to retrieve data and damage the contents of the website, there are several ways, one of which is the BruteForce hacking technique. This technique with the concept of taking, hacking passwords from user administrators on a website, so they can access into the destination website with the level as an administrator [3]. The meaning of BruteForce itself is to force roughly, so it can be interpreted that BruteForce is a technique for hacking a password (password cracking) by trying all possible combinations that exist in the "wordlist" file containing user and password information, which is then synchronized with data user and password of the destination website, or the website that becomes the destination of the attack, forcibly [4]. The logic in practice that directs attacks with this BruteForce technique, is not as easy as imagined, because it requires quite a long time in synchronizing, between the "wordlist" file and the original user data / password. There is a string matching step that requires considerable time. BruteForce algorithm formulation also appeared, which is often used by programmers in making small applications such as password cracking applications, one of which is used in this research, called Hydra. So the BruteForce hacking technique is categorized as a dangerous technique, or a dangerous attack on the website [5].

Browsing data on a network such as the internet, it is not easy, because anyone can access the internet, without the knowledge of the website owner. Because of this, a website administrator is required to understand and protect his website from being attacked by hackers, especially hacking with BruteForce techniques. This search can be done through the classification of data that passes through the network by the Naïve Bayes method, where data passes through the network, the data traffic (traffic) or people accessing the website, it is doing good, only accessing and using it properly, or trying to attack. The point between normal / normal traffic and attack traffic on an internet network can be distinguished. This classification or the separation between normal / normal traffic and attack traffic can be used as a reference, both the amount, speed, or the type of data itself. Assisted by a number of other small tools and applications, it is certain that there are differences, whether it is data that passes through the network containing a BruteForce attack, or whether it is not an attack [6].

The results obtained by using the Wireshak application in the Hydra process which carries the "wordlist" file by doing telnet service, can clearly be distinguished by seeing the traffic display through Wireshark, between normal / normal traffic and the traffic that has attacks. This is also the goal of research. In the notes listed 0.9997 seconds minus 0.9992 seconds between the passage or entry of the normal telnet service process, meaning that without carrying a "wordlist" file with a telnet service that carries a "wordlist" file. Even if compared to the percentages between normal browsing activities or normal internet surfing, but while attacking, that is by doing telnet and Hydra services carrying "wordlist" files is 50% attack traffic, 12.5% normal / normal traffic, at one time.

From several incidents or attacks with the BruteForce technique, it allows more experience to prevent it. The processing and maintenance of a website is getting stronger, especially against attacks with BruteForce techniques. Automatically the benefits of this experience and knowledge can also be used by others. Prevention in question is like locking a hacked account, authentication on the website cookies device, so that it can be filtered, between real users and fake users, and prevention by using Captha [7,10]. Websites that provide accurate, correct and safe data are the hope of users and managers.

2. Research Methods

The method used in this study is Naïve Bayes. With this method can classify and differentiate into categories, between normal traffic and attack traffic. To complement the activities in the study also used several tools or supporting applications,including:

Tabel 1.
Research Support Tools

System	Tools	Information
www.ptrdp.co.id telnet Ubuntu Server 16.04	WireShark Hydra Linux	Visualisasi Service Operating System

In table 1, it is explained that in the Hydra process with telnet service, and in the network detection using the WireShark tools, it will be seen clearly, visually or in GUI and very easily understood by others, which ones are attack traffic and which ones are normal / normal traffic . By using the Naïve Bayes

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

method, it will be calculated based on the time the data passes through the network. Time difference, as said in the Naïve Bayes theory, that a sample will be known next because there is similarity with the previous sample. This means that suppose that if a data passes through a network, then we can see the data that passed through the network before. If the data is similar, both in terms of size, speed and port used, then the data can be determined, normal or attack. That is, if the data were previously normal, then the next one is the same as him, both from the size, speed, and port used, then we can be sure the data is normal. But on the contrary, if the data previously indicated attack data, then come back the next data with the size, speed and port used in common, then certainly the data is attack data.

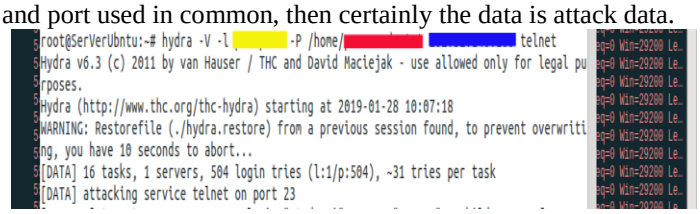

Fig 1. The command line runs the Hydra process with the telnet service on linux Ubuntu Server 16.04

In Figure 1, it appears that with the command line in linux, the hydra command together with the telnet service can be done in one line, or one command at a time. In Figure 1, the yellow box is the user administrator, the red box is the "wordlist" file containing user data and password information that is included or that will be taken by Hydra and telnet when entering the computer network, and then the blue box is ip address being attacked / destination. In Figure 1, it can be explained that the data carried in the Hydra process and telnet service is in the form of a file containing random user and password information. Please note, that Hydra is one of the tools to crack passwords which is included in the hacking category with the BruteForce technique. Telnet service is the act of accessing a computer remotely (remote). The telnet function itself can access easily, such as computer login to the internet and various services remotely. [8]

The following formula is used to identify evidence of data traffic on the network, especially data traffic on the TCP / IP protocol in a computer network.

$$P(\omega \mid x) = P(\omega \mid x1, x2, \ldots, xn) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

With the Bayes theorem, we get:

$$P(\omega \mid x1, x2, \ldots, xn$$

$$(P(\omega) p(x1, x2, \ldots, xn \mid \omega) P(\omega \mid x1, x2, \ldots, xn) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Assuming Naive Bayes that each parameter is independent of the other parameters, equation 1 becomes:

$$P(\omega \mid x) = 1 / C.P(\omega) \quad \text{atau} \quad \Pi^{n}{}_{i} = 1.p(xn \mid \omega) \ldots\ldots\ldots (3)$$

Where $\boldsymbol{C = p(x1, x2, \ldots, xn)}$ **is a scale factor.**

Where C (proof) is a scale factor depending only on x1, x2, x3 ..... Xn is a constant if the value of the feature variable is known. [6]

In this study, data traffic on the network measured is data that passes through the TCP / IP protocol. This does not mean ignoring other ports through the protocol, but it can be easier to focus on just one port, so that data that passes through the network or its traffic will be detected properly and can also be easily classified by the Naïve Bayes method, because in Naïve Bayes theory , conceptualizing sorting based on the same category. The formula of Naïve Bayes which is listed up to 3 steps, can be assumed that one formula will be able and succeed the same as the other formula. The same result will be obtained, with only one formula. The formulation is based on the variables used in the study. Here what is used is the formulation of step 3, because it strongly represents the variables used in the study. Variables referred to in this study are all matters related to the research process and are included in the research process, as mentioned in the formula "P" is a parameter, "C" is Evidence, "X" is a constant, which includes frame traffic when Hydra processes and telnet services are run into a computer network, "ω" is Posterior Propability, which is Prior Probability which is a probability value that someone believes to be right before experimenting on something. If the experiment is then carried out and results in a change or improvement to the probability value, then it is referred to as Posterior Probability. This is in accordance with the Naïve Bayes algorithm, which is to predict future opportunities based on past experiences, or based on past events that have been missed, so that in this study the Naïve Bayes formula is used to classify per frame on PCAP (Pattern Capture) displayed by the Wireshark application with a count of

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

seconds per frame, between the first frame and the next frame. If there are similarities between one frame and the other, it is certain that it is the same process. The description here is that the time between frames on traffic will be calculated and then classified, whether it is normal / normal traffic, or attack traffic. [9] [google searching]

3. Result

3.1 Normal Traffic

Normal traffic in question is where the data traffic that passes through or through the TCP / IP protocol is the activity of browsing, downloading and streaming video as in general. Detection is carried out using WireShark tools whose function is to retrieve realtime normal / normal traffic at the time of the study. PCAP (Pattern Capture) of existing / normal traffic will display in detail with a calculation using Naïve Bayes.

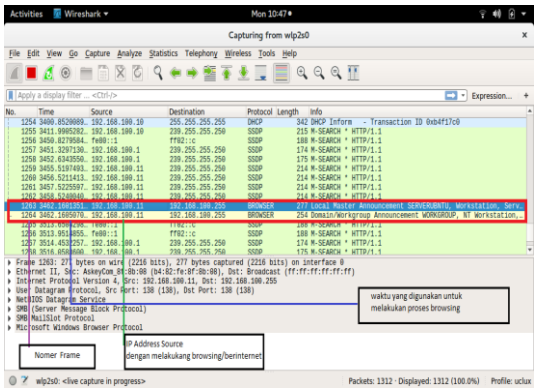

Fig 2. Detect with WireShark on normal / normal traffic

The process in Figure 2 was carried out on January 28, 2019 with the source IP address 192.168.100.11 with IP destionation 192.168.100.255 in the process of browsing, surfing. Analyzing using theorems, the Naïve Bayes formula is obtained as follows:

Known :

X = Data sample / Frame
X = X1 at frame 1263, X2 at frame 1264
ω = Posterior propability
P = Parameter
Parameter = Browser process, there are 2 frames x1 and x2
C = Evidence value / Scale factor

Calculated by the time variable required to perform two processes, the analysis results are:

Frame 1263 with a time of 3462,1601 seconds while frame 1264 with a time of 3462,1605 seconds, so that the time interval required for the browser process by IP 192.168.100.11 is equal to the time on frame X2 reduced by the time on frame X1 which is 0,0004 seconds.

The formula X2-X1 = 3462,1605 - 3462,1601 = 0,0004

Then the value of C as a scale factor is:

C = P (X1, X2)
= 1. (0,0004)
= 0,0004 seconds

This means that the appearance of the browser process between frame X1 and X2 takes 0,0004 seconds. The next step is to calculate the posterior propability with the following formula:

$$P(\omega|x) = 1/C \,.P(\omega)$$
$$= 1/0{,}0004 \,.1.(\omega)$$
$$\omega = 1/0{,}0004.\ ½$$
$$= 1/0{,}0008$$

This illustrates that the posterior propability of ordinary / normal and general traffic (browsing) will likely

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

occur at an average time of 1 / 0.0008 seconds for every 1 second. This means that within 100% of an event the following event will probably occur with the following time:

$$(100/100) \,.\, (1/0{,}0008) = 100/0{,}08 = 1250$$

$$1{:}0{,}08 = 12{,}5 \quad =====\rightarrow 12{,}5\%$$

Get a 12.5% second result for the appearance of the same process in 100% or 1 second.

3.2 Normal / Normal Telnet Traffic Service

This means that when running a telnet service, it does not carry data files for cracking passwords or does not intend to run a BruteForce attack, only activity on a normal telnet traffic service is directed to a particular website address. IP source 192.168.100.11 does telnet service to IP 202.52.146.106 (www.ptrdp.co.id). Telnet service, realtime and does not carry data, meaning that it is pure telnet only to the website. In this service, the traffic seen on the Wireshark tool is data traffic from the telnet service that goes to the www.ptrdp.co.id website through the TCP / IP protocol. Display in the Wireshark tool, transmission occurs repeatedly at a time. In this case the Naïve Bayes theorem calculates data traffic only on the TCP / IP protocol.

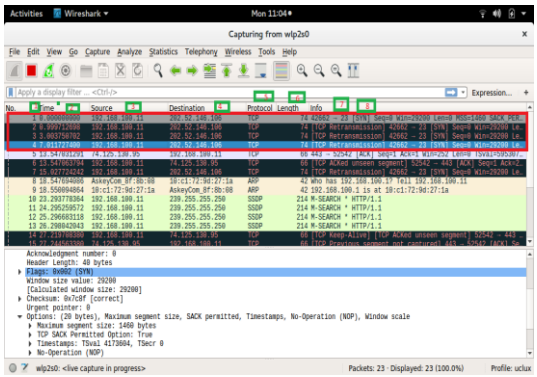

Fig 3. WireShark detection on telnet services without carrying data

Seen from left to right

1. Frame Number
2. Processing time of service
3. IP Source
4. Destination IP
5. Protocol
6. Data length
7. Port Info
8. Telnet Service Info

Indication that Figure 3 shows the telnet service is via TCP protocol with port 23. Between one frame with the next frame with the same data length that is 74 bytes. The normal telnet service process is then calculated using the Naïve Bayes theorem formula as follows:

Known :

X = Data sample / Frame

X = X1 at frame 1, X2 at frame 2

ω = posterior propability

P = Parameter

Parameter = normal / normal telnet service, there are 2 frame, x1 (mass synchronize) and x2 (mass process)

C = Evidence value / Scale factor

Calculated by the variable time needed to perform normal telnet services between the initial entry time with the synchronization process and then in the second frame starts transmitting data from the telnet itself, the following analysis results are obtained:

The distance, interval or time range needed by both processes in frame 1 and frame 2 is recorded, at frame 1 or X1 = 0.0000 seconds. Then in frame 2 which is stated by X2 = 0.9997 seconds. The result between the two frames is calculated by the value of the variable interval between X1 to X2.

X1 = 0.0000 seconds X2 - X1 seconds

X2 = 0.9997 0.9997 - 0.0000 = 0.9997 seconds

The value of C or indicative evidence as a scale factor with the Naïve Bayes theorem formula is:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

C = P (X1, X2)
= 1. (0.9997)
= 0.9997 seconds

This means that the time required to perform normal telnet service / normal which is a parameter of Naïve Bayes propability, between frame 1 to frame 2 is 0.9997 seconds.

The results obtained by calculating through the Naïve Bayes theorem formula are as follows:

P (ω | x) = 1 / C .P (ω)
= 1 / 0.9997 .1 (ω)
ω = 1 / 0.9997. ½
= 1 / 1.9994

This means that the posterior propability of a normal / normal telnet service will emerge from the synchronization period to the transmission of data requiring an average time of 1 / 1.9994 seconds for every 1 second.

The assumption is that within 100% of an event an event might reappear over time
(100/100). (1 / 1.9994) = 100 / 199.94 = 0.5001 = 50.01
1: 1.9994 = 0.5001 seconds ===➔ 50%

The results state that within 1 second the normal / normal telnet service process will emerge from synchronization to the average data transmission takes 0.5 seconds or mentioned with 100% of the possibility that there is a 50% normal / normal telnet traffic service will occur again.

3.3 Traffic BruteForce, Telnet Service By Bringing Data to the Hydra Process

In this session the telnet service is accompanied by a Hydra process, meaning that traffic or data traffic through the TCP / IP protocol is a telnet service carrying BruteForce attack data that is the Hydra process, because in the Hydra process, the data carried is in the form of user data lists and the password. The traffic displayed by the Wireshark tool is a hack / attack technique with BruteForce.

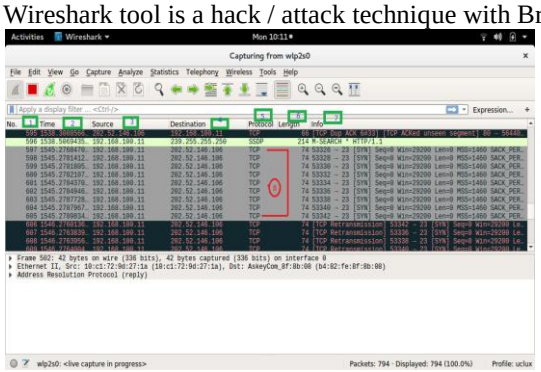

Fig 4. WireShark, Telnet Service and Hydra Process

The description of Figure 4 is:

Seen from left to right

a. Frame Number 5. Protocol
b. Processing time of service 6. Data length
4. IP Source 7. Port Info, Process
5. Destination IP 8. Synchronization Period

It is indicated that in Figure 4 the process is recorded on port 23 with the same data length of 74 bytes. The protocol used by the service based on Figure 4 is fixed on TCP so that it can be ensured that the service that occurs in accordance with the existing template, is a telnet service even though it is combined with the Hydra process. The results obtained by calculating the Naïve Bayes formula as below.

Known :

X = Data sample / Frame
X = X1 at frame 597-605, X2 at frame 606
ω = posterior propability
P = Parameter
Parameter = Service telnet with Hydra, there are 10 frame x1 (mass synchronize) and x2 (mass process)
C = Evidence value / Scale factor

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

The time required by the telnet service together with the Hydra process with two frames categorized between X1 to X2 in the number of frames of 10 sessions is to calculate the time interval. Reduction of the time recorded in X2 - X1 (first frame) so that there is a formula for the time difference, the time interval between the two.

X1 = 1545.2768 seconds
X2 = 1546.2760 seconds
X2 - X1 = 1546.2760 - 1545.2768 = 0.9992 seconds

Maka nilai dari Cadangan bukti atau faktor skala dihasilkan dengan rumus

$$C = P(X1,X2)$$
$$= 1 \,.\, (0{,}9992)$$
$$\mathbf{= 0{,}9992 \text{ seconds}}$$

Described as the time required for the Telnet service coupled with the Hydra process of X1 frame synchronization for data transmission frames equals 0.9992 seconds. Then the calculated equation from Naïve Bayes is:

$$P(\omega \mid x) = 1/C \,.P(\omega)$$
$$= 1/0{,}9992 \,.1.(\omega)$$
$$\omega = 1/0{,}9992.\ ½$$
$$= 1/1{,}9984$$

It is assumed that posterior propability that a telnet service along with the Hydra process will reappear from synchronization to data transmission requires an average time of 1 / 1.9984 seconds from every 1 second. This means that within 100% of a telnet service event coupled with the Hydra process will reappear in the assumed time size of:

(100/100). (1 / 1.9984) = 100 / 199.84 = 0.5004 = 50.04

1: 1.9984 = 0.5004 seconds ===➔ 50%

The assumption is that within 100% of the telnet service will appear coupled with the Hydra process with an average appearance time of 50% seconds. When compared with the normal telnet service process, the resulting time count is only 0,0003 seconds faster. It can be mentioned that the normal telnet service / traffic appears, 0,0003 seconds faster than the telnet service coupled with the Hydra process.

4. Conclusion

In the hacking process there are several techniques, one of which is the BruteForce technique. This technique is very simple, but arguably it is deadly. This is because in the BruteForce technique there is a process of matching users and passwords of a website administrator who is often known as the password crack application. If a hacker can access to other people's websites, it can also retrieve, copy, edit and add data and even damage it. However, this BruteForce technique can be tracked earlier so that it can be avoided or prevented. This research is on TCP / IP protocol with port 23 of the traffic / data traffic that passes through it. The results can be distinguished and characterized by general, normal, normal traffic such as browsing activities on the internet, normal / normal telnet services, and telnet services by bringing the BruteForce process, namely Hydra. Each has the results of calculations with the Naïve Bayes theorem formulation as a tool for classification, namely that the act of browsing on the internet network requires 0,0004 / sec or 12.5% / sec while for attack data using telnet services coupled with the hydra process (BruteForce technique)) takes 0.9992 / second or 50% / second. An accuracy rate of 12.5% for the browsing process and 50% for the BruteForce technique, meaning that attacks appear to be greater than browsing on the internet, this means that if data traffic occurs the same thing on an ongoing basis, it can be said and characterized that it is possible to attack data traffic. Prevent with some techniques, can protect the user database and password itself, such as locking an account, using Captha, authentication device cookies. Other ways such as using a combination of characters that are difficult on the password, and or using a sophisticated firewall on the server, so that hackers can not enter the network.

5. References

[1] Y. Pratomo, "Hacker Klaim Jual 26 Juta Akun Internet." KompasOnline, p. https://tekno.kompas.com/read/2019/03/18/11250077/, 2019.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Jurnal Mantik
Volume 4 Number 1 May 2020, pp. 319-326 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

[2] R. Umar, A. Yudhana, and M. N. Faiz, "Experimental Analysis of Web Browser Sessions Using Live Forensics Method Experimental Analysis of Web Browser Sessions Using Live Forensics Method," no. December, pp. 2951–2958, 2018.
[3] Feradhita, "Mengenal Brute Force Attack dan Cara Menghindarinya." 2019.
[4] S. Sandra, D. Stiawan, and A. Heryanto, "Visualisasi Serangan Brute Force Menggunakan Metode K-Means dan Naïve Bayes," Proceeding - Annu. Res. Semin. Proceeding, vol. 2, no. 1, pp. 315–320, 2016.
[5] E. M. . Fenty, A. Hanifa, and N. Riyanto, "Implementasi Algoritma Brute Force dan Fitur Location Based Services (LBS) Pada Aplikasi Kumpulan Doa Harian Berbasis Android," vol. 2. p. 11, 2014.
[6] L. S. Malang et al., "Penerapan Metode Naive Bayes dalam Pengklasifikasi Trafik Jaringan," Smatik J., vol. Vol 06, no. January 2016, pp. 26–36, 2017.
[7] P. Andalas, "Mengenal Teknik Hacking Brute Force," 2017.
[8] H. Pramaditya, "Brute Force Password Cracking Dengan Menggunakan Graphic Processing Power," J. Teknol. dan Manaj. Inform., vol. 2, no. 1, 2016.
[9] M Ferdy Adriant dan Is Mardianto, "Implementasi Wireshark Untuk Penyadapan (Sniffing) Paket Data Jaringan," Semin. Nas. Cendekiawan, pp. 224–228, 2015.
[10] Ritzkal, Keamanan Jaringan Cyber,,Bogor:UIKA PRESS,2019

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.