Analisis Tren Lalu-lintas Data Jaringan Menggunakan Teknologi Big Data (Studi Kasus: UNIVERSITAS MAHADEWA INDONESIA)

Gemara Adhiyasa Parahita Nugraha[a1], I Made Agus Dwi Suarjaya[a2], I Putu Agus Eka Pratama[a3]
[a]Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana
e-mail: [1]gema.sesetan@gmail.com, [2]agussuarjaya@it.ac.id [3]eka.pratama@unud.ac.id

Abstrak

Universitas Mahadewa Indonesia Merupakan lembaga pendidikan pembelajaran yang memiliki jaringan Wifi sebagai penyedia internet kepada mahasiswa dan dosen di Universitas Mahadewa. Jaringan Wi-fi tentunya banyak di akses oleh mahasiswa maupun pegawai di Universitas Mahadewa. *Big Data* merupakan suatu *trending practice* yang banyak dilakukan oleh berbagai perusahaan, organisasi, peneliti, maupun akademisi yang bertujuan untuk untuk mendapatkan suatu informasi yang penting dalam data tersebut. Penelitian ini membahas evaluasi data yang di dapat dari *monitoring* jaringan di Universitas Mahadewa. Wireshark atau Tshark adalah *tools monitoring* yang di pergunakan untuk monitoring jaringan pada suatu *server* yang mampu menangkap data pada jaringan wifi yang di akses oleh *user* tersebut. Salah satu pengimplementasian *Big Data* yaitu pada Hasil dari penangkapan *packet data* tersebut akan di analisis untuk di lihat hasilnya sehingga penulis mengetahui penggunaan *Domain* di jaringan Universitas Mahadewa. Berdasarkan pengujian *monitoring* jaringan menggunakan Wireshark atau Tshark, didapat hasil *packet data user* yang mengakses website menggunakan jaringan Universitas Mahadewa. Tujuan penelitian ini untuk mengetahui penggunaan Wireshark atau tshark dan analisi data dalam jumlah banyak dari hasil *monitoring* tersebut. Metode untuk analisis data ini adalah visualisasi dimana mahasiswa menggunakan aplikasi Tableu untuk visualisasikan data dalam jumlah jutaan atau banyak.

Kata kunci: Jaringan(Wi-fi), Capture Data, Wireshark, Packet data, Big Data, Tableu.

Abstract

Indonesia Mahadewa University is a learning education institution that has a Wi-fi network as an internet provider for students and lecturers at Mahadewa University. Wifi network is of course widely accessed by students and employees at Mahadewa University. Big Data is a trending practice that is mostly carried out by various companies, organizations, researchers, and academics with the aim of obtaining important information in the data. This study discusses the evaluation of data obtained from network monitoring at Mahadewa University. Wireshark or tshark is a monitoring tool that is used for network monitoring on a server that is able to capture data on a wi-fi network that is accessed by the user. One of the implementations of Big Data is that the results of capturing packet data will be analyzed to see the results so that the authors know the use of domains in the Mahadewa university network. Based on network monitoring testing using Wireshark or Tshark, the results obtained are packet data users who access the website using the Mahadewa University network. The purpose of this study is to determine the use of Wireshark or Tshark and analyze large amounts of data from the monitoring results. The method for data analysis is visualization where students use the Tableu application to visualize data in millions or lots.

Keywords: Network (Wi-fi), Capture Data, Wireshark, Packet data, Big Data, Tableu.

1. Introduction

Beberapa penelitian sebelumnya yang dijadikan acuan dalam pengerjaan Tugas Akhir dengan judul Manajemen Keamanan Jaringan Berbasis Wireshark untuk jaringan Sehat di perusahaan akan dipaparkan sebagai berikut.

Penelitian oleh Junaidi yang berjudul deteksi serangan pada jaringan komputer dengan Wireshark menggunakan metode *Anomally-Based* IDS yaitu melakukan implementasi sistem pendeteksi serangan jaringan komputer menggunakan metode Packet Streaming, dimana pemasangan aplikasi Wireshark pada komputer yang memakai sistem operasi Windows untuk melakukan monitoring jaringan. [1]

Penelitian oleh Edwin yang membahas tentang analisis keamanan jaringan internet(wi-fi) dari serangan *packet data Sniffing.* Untuk memahami sistem *forensic* pada suatu jaringan, kegunaan beberapa peralatan yang berhubungan dengan sistem jaringan, sehingga kita mengetahui sistem kinerja peralatan-peralatan tersebut. Penggunaan ini juga sangat berguna ketika kita melakukan analisis dan investigasi paket data jaringan. [2]

Penelitian oleh Didi Susianto, dan Anisa Rachmawati mengenai implementasi dan analisis jaringan menggunakan Wireshark, Cain and Abels, Network Miner membahas tentang melakukan pengumpulan data. Metode yang digunakan dalam penelitian ini adalah studi kasus, menggunakan cara-cara yang sistematis dalam melakukan pengamatan, pengumpulan data, analisis informasi, dan pelaporan hasilnya. Perancangan Sistem Teknik PGP dalam melakukan monitoring ini, peneliti masuk ke jaringan Wifi yang tersedia pada AMIK Dian Cipta Cendikia Bandar Lampung Khususnya di kampus A yang dapat digunakan mahasiswa AMIK Dian Cipta Cendikia Bandar Lampung. Sebelum melakukan monitoring, Software Cain and Abel dijalankan terlebih dahulu untuk melakukan *routing* agar *traffic* jaringan bisa terpantau. [3]

Penelitian oleh Russ McRee yang berjudul *Security Analysis* with Wireshark melakukan analisis perilaku mallware mengumpulkan data VMnet8 untuk menangkap lalu lintas dari OS yang terinfeksi dengan analisis TCP Stream dan di analisis untuk di ketahui *malware* menyerang protokol apa saja dan bisa dilakukan penanganan atau pencegahan pada *malware* tersebut. [4]

2. Research Method / Proposed Method

Metodologi penelitian yang digunakan dalam tugas akhir ini terdiri dari empat tahap, yaitu observasi, studi literatur, perancangan sistem, dan penelitian. empat tahap tersebut akan dijelaskan pada bagian berikut.

2.1. Observasi

Analisis merupakan langkah awal dalam peneilitian ini. Data yang akan digunakan diambil melakukan capture Data Jaringan menggunakan Wireshark. Wireshark diinstall di komputer server untuk lakukan penelitian dan ditinggal selama 4 bulan agar mendapat data yang akan dianalisis. Data yang sudah dicapture akan disimpan dalam *format* csv untuk dibersihkan agar dapat dianalisis ketahap selanjutnya.

2.2. Studi Literatur

Studi literatur dilakukan dengan memfokuskan pencarian teori-teori penelitian yang akan digunakan untuk penelitian, dokumentasi dari teknologi yang dipakai, serta fitur-fitur dari tools yang digunakan dalam penelitian ini. Literatur yang digunakan terkait dengan topik penelitian yaitu terutama mengenai Capture Data. Kemudian dalam Analisis Data dibutuhkan studi literatur mengenai teknologi Big Data untuk Analisis Data Capture.

2.3. Perancangan Sistem

Tahapan ini melakukan perancangan sistem yang akan dibuat dan diujikan, yang meliputi perancangan penelitian Analisis Tren Lalu-lintas Data Jaringan Menggunakan Teknologi Big Data. Wireshark adalah *tools* yang digunakan optimalisasi Jaringan Internet untuk manajemen jaringan internet di Universitas Mahadewa.

2.4. Pengujian Sistem

Penelitian ini dilakukan dengan melakukan *bridge* jaringan pada switch distributor, switch access, dan router distributor diarahkan ke computer server untuk melakukan penelitian *capture* data jaringan, fungsi dari *bridge* jaringan *server* adalah untuk menjadikan computer server sebagai man of the middle sehingga segala aktivitas akan melewati *computer server* dan capture dengan Wireshark atau Tshark sehingga peneliti akan mengetahui aktivitas *user* yang menggunakan wifi Universitas Mahadewa Indonesia.

3. Literature Study

Teori yang digunakan dalam penelitian tentang Analisis Tren Lalu-lintas Data Jaringan Menggunakan Teknologi Big Data adalah sebagai berikut.

3.1. MongoDB Database Storage

MongoDB adalah salah satu jenis database NoSQL yang cukup populer digunakan dalam pengembangan *website*. Berbeda dengan database jenis SQL yang menyimpan data menggunakan relasi tabel, MongoDB menggunakan dokumen dengan format JSON dan CSV. Hal inilah yang dianggap membuat pengelolaan data menggunakan MongoDB lebih baik. [5]

3.2. Tools Wireshark

Wireshark adalah *tools* yang ditujukan untuk penganalisaan paket data jaringan. Wireshark disebut juga *Network packet analyzer* yang berfungsi menangkap paket-paket jaringan dan berusaha untuk menampilkan semua informasi dipaket tersebut sedetail mungkin. Sebenarnya *network packet analyzer* sebagai alat untuk memeriksa apa yang sebenarnya terjadi di dalam jaringan baik kabel maupun wireless. Dengan adanya Wireshark ini semua sangat dimudahkan dalam hal monitoring dan menganalisa paket yang lewat di jaringan. [6]

3.3. Big Data

Big Data adalah istilah umum untuk segala kumpulan himpunan data dalam jumlah yang sangat besar dan kompleks sehingga menjadikannya sulit untuk ditangani atau di proses jika hanya menggunakan manajemen basis data biasa atau aplikasi pemroses data tradisional. Big Data menjamin pemrosesan solusi data dengan varian baru maupun yang sudah ada untuk memberikan manfaat nyata bagi bisnis. Namun pengolahan data dengan ukuran dan kompleksitas besar tetap sekedar solusi teknologi kecuali jika dikaitkan dengan tujuan bisnis. Big Data mempunyai potensi yang sangat besar dalam pengambilan keputusan dan menjadi suatu solusi hampir disetiap organisasi untuk membantu memprediksi pola perilaku pembelian pelanggan, mendeteksi penipuan, dan penyalahgunaan. [7]

3.4. Tableu

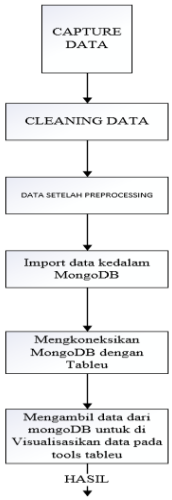
Perancangan visualisasi data merupakan serangkaian tahapan dalam menampilkan hasil analisis data menggunakan tools Tableau. Tableau merupakan *software* yang dirancang untuk membuat visualisasi data, laporan, dan dasbord dengan cara yang cepat. Tableau dapat terhubung ke beberapa sumber data, melakukan analisis data multidimensi, membuat dasbor atau laporan. Tableau merupakan pemrosesan data yang efisien dan mempunyai *user interface* yang mudah digunakan (Batt et al., 2020). Visualisasi data menggunakan tools Tableau yang dimulai dari *import* data dari *database* MongoDB, kemudian dilanjutkan dengan proses menambahkan *measure latitude* dan *longitude* untuk menentukan titik koordinat-nya, selanjutnya memilih filter *dimensions* berupa waktu, Info, dan yang terakhir menampilkan hasil Hitungan terbanyak akses domain pada Universitas Mahadewa. [8]

4. Result and Discussion

Hasil dan diskusi pada penelitian Analisis Tren Lalu-lintas Data Jaringan Menggunakan Teknologi Big Data (Studi Kasus Universitas Mahadewa Indonesia) membahas gambaran umum sistem, hasil *Capture Data*, hasil menyimpan data pada Mongodb, serta hasil visualisasi data.

4.1. Gambaran Umum Sistem

Gambaran umum sistem dibuat dalam bentuk topologi jaringan, proses metode secara umum ini mendeskripsikan mengenai analisa proses dalam penelitian secara umum dan dalam analisa proses metode secara umum ini ada dua proses yaitu proses pengambilan data, proses analisa data, dan visualisasi data. Berikut merupakan tampilan dari gambaran umum Analisis Tren Lalu-lintas Data Jaringan Menggunakan Teknologi Big Data.



Gambar 1. Gambaran Umum Sistem.

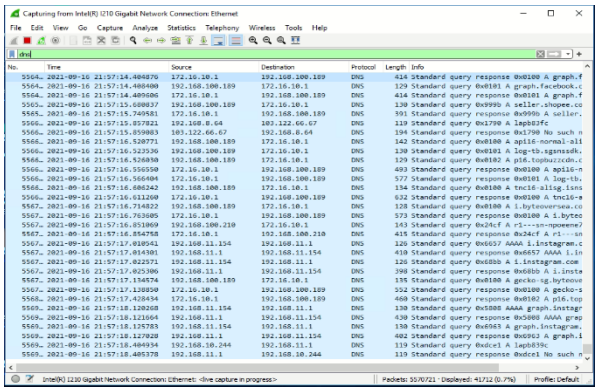
Gambar 1 merupakan gambaran umum sistem dalam bentuk desain jaringan. Wireshark atau Tshark akan bertindak sebagai *software* yang dapat melakukan *monitoring* terhadap jaringan di Universitas Mahadewa. *Router* sebagai penghubung ke dalam internet dan domain. Visualisasi data dengan *database* MongoDB yang terintegrasi dengan Tableau membuat visualisasi data menjadi lebih cepat bahkan dengan jumlah data yang besar.

4.2. Pengumpulan Data Capture Jaringan

Pengumpulan data *capture* merupakan tahapan pengumpulan data yang digunakan sebagai data uji. Data *capture* ini berupa DNS dikumpulkan menggunakan filter DNS dari Tshark. Struktur data *capture* ini berupa format csv yang kemudian disimpan pada database MongoDB. Detail mengenai struktur data *capture* dapat dilihat pada Gambar 2 berikut.

Gambar 2. Proses *Capture* DNS Tshark.

Gambar 2 memperlihatkan proses *capture* dari Tshark berupa format csv yang disimpan pada *database* MongoDB untuk di*preprocesing* dilanjutkan visualisasi data. Berikutnya pengumpulan data *capture* versi Wireshark dapat dilihat pada gambar 3.



Gambar 3. Proses Capture DNS Wireshark.

Gambar 3 memperlihatkan proses *capture* dari Wireshark berupa format csv yang disimpan pada *database* MongoDB untuk di*preprocesing* dilanjutkan visualisasi data. Berikutnya hasil data *capture* dapat dilihat pada gambar berikut.

Gambar 4. Hasil Pengumpulan Data Capture Jaringan.

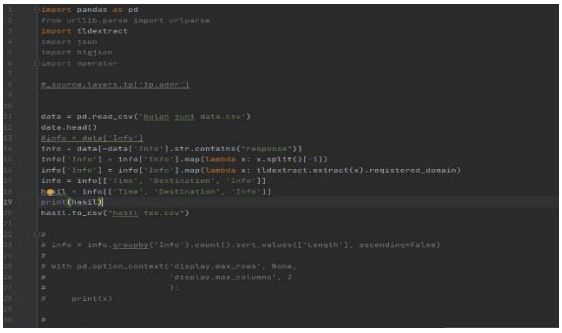
Gambar 4 menjelaskan mengenai hasil proses pengumpulan data capture jaringan dari jaringan Universitas Mahadewa menggunakan Wireshark dan Tshark dengan mengumpulkan seluruh *protocol* DNS secara lengkap.

4.3. Hasil Analisa Data

Hasil Analisa data merupakan hasil dari serangkaian tahapan *pre-processing* data dan *processing* data. Tahap *pre-processing* data yaitu tahap yang dimana proses membersihkan data mulai dari proses *cleaning* dimana data dibersihkan menggunakan Bahasa pemograman python untuk diambil bagian yang penting agar bisa di analisis.

4.3.1. Hasil Proses Pembersihan Data

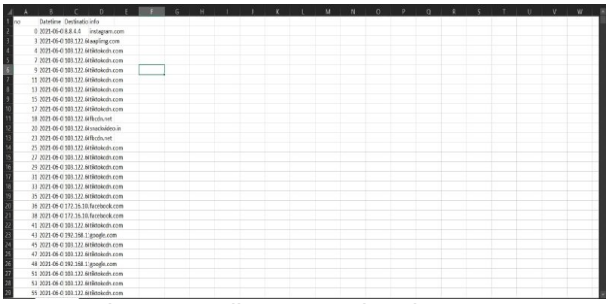
Proses pembersihan data merupakan tahapan dimana data dibersihkan dari text yang tidak di perlukan sehingga mempermudah analisis data sehingga bisa di lanjutkan ke tahap visualisasi. Detail mengenai struktur pembersihan data dapat dilihat pada Gambar 5.



```
import pandas as pd
from urllib.parse import urlparse
import tldextract
import json
import ndjson
import operator

#_source.layers.tp['ip.addr']


data = pd.read_csv('bulan juni data.csv')
data.head()
#info = data['Info']
info = data[~data['Info'].str.contains("response")]
info['Info'] = info['Info'].map(lambda x: x.split()[-1])
info['Info'] = info['Info'].map(lambda x: tldextract.extract(x).registered_domain)
info = info[['Time', 'Destination', 'Info']]
hasil = info[['Time', 'Destination', 'Info']]
print(hasil)
hasil.to_csv("hasil tes.csv")

#
# info = info.groupby('Info').count().sort_values(['Length'], ascending=False)
#
# with pd.option_context('display.max_rows', None,
#                        'display.max_columns', 2
#                        ):
#     print(x)

#
```

Gambar 5. Proses Pembersihan Data.

Gambar 5 memperlihatkan proses pembersihan data dari Bahasa pemograman python berupa format csv yang disimpan pada excel dan di simpan pada MongoDB untuk dilanjutkan visualisasi data. Selanjutnya hasil Proses Cleaning data. Hasil proses *cleaning* data merupakan hasil dari proses pembersihan data dari atribut-atribut *text* yang tidak digunakan dalam analisis seperti *source*, *length*, dan nomer. Hasil proses *cleaning* data dapat dilihat pada Gambar 5.

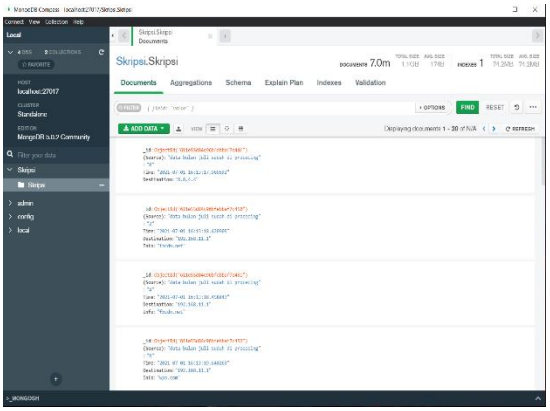


Gambar 6. Hasil Proses *Cleaning* Data.

Gambar 6 menampilkan mengenai hasil dari proses *cleaning* data *capture* Jaringan yaitu menghilangkan *source*, *protocol*, dan *length* yang ada pada *text capture* jaringan agar lebih mudah di analisis dengan *tools* Tableu.

4.3.2. Hasil Proses Penyimpanan

Hasil penyimpanan data *capture* jaringan *clean* ini merupakan penyimpan data dari hasil *pre-processing* yang dilakukan pada *text capture*, kemudian disimpan pada database MongoDB. Hasil penyimpanan data *capture* jaringan *clean* dapat dilihat pada Gambar 4.4 berikut.

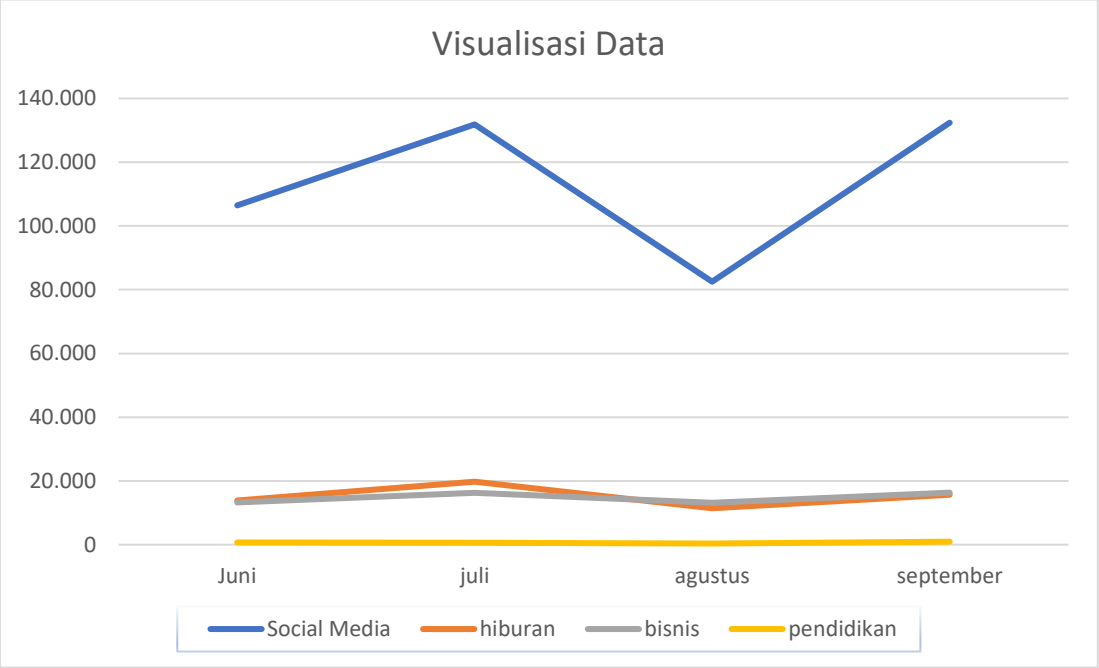


Gambar 7. Hasil Penyimpanan Data *Capture* Di Mongodb.

Gambar 7 menjelaskan mengenai hasil penyimpanan data *capture* jaringan yang disimpan dalam format CSV (*Comma Separated Values*) pada *database* MongoDB.

4.4. Visualisasi Data

Visualisasi data merupakan hasil visualisasi data analisis Domain berupa bentuk grafik *line* dan memiliki kategori di setiap domain terhitung Juni tahun 2021 sebelum PPKM di mulai sampai dengan dengan bulan september tahun 2021 setelah PPKM di mulai selama kondisi pandemi COVID-19. Hasil mengenai bentuk visualisasi data memiliki beberapa kategori dapat dilihat sebagai berikut.



Gambar 8. Hasil Visualisasi Data 4 bulan.

Gambar 8 menjelaskan mengenai hasil visualisasi data selama 4 bulan. Dalam kategori sosial media di Bulan Juni terdapat 106.425 total akses, di Bulan July terdapat 131.824 total akses, di Bulan Agustus terdapat 82.508 total akses, dan di Bulan September terdapat 132.394 total akses. Dalam kategori hiburan di Bulan Juni terdapat 13.899 total akses, di Bulan Juli terdapat 19.752 total akses, di Bulan Agustus terdapat 11.422 total akses, dan di Bulan September terdapat 15.634 total akses. Dalam kategori bisnis di Bulan Juni terdapat 13.208 total akses, di Bulan Juli terdapat 16.269 total akses, di Bulan Agustus terdapat 13.147 total akses, dan di Bulan September terdapat 16.316 total akses. Dalam kategori pendidikan di Bulan Juni terdapat 723 total akses, di Bulan Juli terdapat 650 total akses, di Bulan Agustus terdapat 361 total akses, dan di Bulan September terdapat 944 total akses. Pada visualisasi di atas bahwa kategori sosial media banyak yang mengakses ketimbang kategori pendidikan, hiburan, dan berita. Dengan kata lain *user* di Universitas Mahadewa lebih sering mengakses sosial media.

5. Conclusion

Wireshark dan Tshark dapat dimanfaatkan sebagai *tools* untuk *monitoring* jaringan melalui *server* dengan tujuan mengetahui apakah jaringan di Universitas Mahadewa efektif melakukan pembelajaran contoh berapa banyak *website* pendidikan di akses oleh mahasiswa atau dosen, teknologi *big data* dapat dimanfaatkan untuk mengolah dan menganalisis data yang banyak, tujuan kedua menggunakan teknologi *big data* yaitu mengetahui *trend* yang berada pada Universitas Mahadewa. *Trend* adalah kebiasaan *user* mengakses *domain* paling tinggi angka aksesnya contoh Instagram jadi divisualisasi akan terlihat dalam beberapa bulan apakah ada peningkatan atau penurunan di setiap bulan.

Daftar Pustaka

[1] C. Z. Xianglin Xiao, "Application and challenges of big data in quality monitoring of highway engineering," in *AIP Conference Proceedings 1820*, Sichuan Chengdu China, 2017.

[2] R. McRee, "Security Analysis with Wireshark," *JOURNAL ISSAC Vol 1 No 5*, p. 1, 2017.

[3] A. L. V. Z. L.U. Laboshin*, "The Big Data approach to collecting and analyzing traffic data," in *Procedia Computer Science*, Polytechnicheskaya street Russia, 2017.

[4] A. M. C. B. Hasan Basri, "ANALISIS KEMANAN JARINGAN INTERNET (WIFI) DARI SERANGAN PACKET DATA SNIFFING," *Jurnal PROSISKO Vol. 4 No 2.*, p. 4, 2017.

[5] E. B. Harjono, "DETEKSI SERANGAN PADA JARINGAN KOMPUTER DENGAN WIRESHARK MENGGUNAKAN METODE ANOMALLY-BASED IDS.," *Jurnal & Penelitian Teknik Informatika Volume 1 Nomor 1.*, p. 3, 2016.

[6] M. M. M. H. F. C. S. Andy Rachman, "IMPLEMENTASI DAN ANALISIS JARINGAN MENGGUNAKAN WIRESHARK, CAIN AND ABELS, NETWORK MINNER," *Jurnal Cendikia Vol.XVI*, p. 2, 2014.

[7] S. N. I. P. Zen Munawar, "Keamanan Jaringan Komputer Pada Era Big Data," *Jurnal Sistem Informasi – J-SIKA Volume 02*, pp. 14-20, 2020.

[8] M. Syafrizal, Pengantar Jaringan Komputer, Yogyakarta: C.V. ANDI OFFSET, 2005.

[9] O. K. Sulaiman, "Analisis Sistem Keamanan Jaringan Dengan Menggunakan Switch Port Security," *CESS (Journal Of Computer Engineering, System And Science)*, pp. 9-14, 2016.

[10] S. C. S. B. Rakesh Kumar, "Effective Way to Handling Big Data Problems using," *STM Journals*, pp. 42-48, 2015.

[11] I. F. S. M. A. B. S. M. M. Junaidi Syahputra, "Deteksi Serangan pada Jaringan Komputer dengan Wireshark Menggunakan metode Anomally Based," *Jurnal & Penelitian Teknik Informatika*, pp. 12-20, 2016.