Volume 11 Issue 2 (2026) Pages 904 - 916
Jurnal Mirai Management
ISSN : 2597- 4084 (Online)

Identifikasi Anomali *Traffic* Data Jaringan Wi-Fi Menggunakan Metode *Isolation Forest* dan *One-Class* SVM

Muhamad Andy ✉ , **Anjik Sukmaaji**[2]
[1,2] *Sistem Informasi, Teknologi dan Informatika, Universitas Dinamika*

Abstrak

Peningkatan penggunaan jaringan Wi-Fi menyebabkan kepadatan traffic data semakin tinggi sehingga diperlukan proses pemantauan untuk mendeteksi aktivitas yang menyimpang dari pola normal jaringan. Penelitian ini bertujuan mengidentifikasi anomali pada data traffic Wi-Fi hasil tangkapan Wireshark tanpa bergantung pada ketersediaan label serangan yang lengkap. Data paket jaringan dikumpulkan menggunakan Wireshark, kemudian diekspor ke dalam format CSV untuk dilakukan preprocessing. Selanjutnya, dibentuk fitur RequestCount sebagai representasi aktivitas permintaan jaringan. Proses deteksi anomali dilakukan menggunakan dua metode unsupervised learning, yaitu Isolation Forest dan One-Class SVM. Isolation Forest bekerja dengan mengisolasi observasi yang memiliki karakteristik berbeda dari mayoritas data, sedangkan One-Class SVM membentuk batas terhadap pola data normal dan mengidentifikasi observasi yang berada di luar batas tersebut sebagai kandidat anomali. Hasil pengolahan berupa data traffic yang telah dibersihkan, nilai RequestCount, anomaly score, serta klasifikasi Normal atau Anomali. Informasi tersebut dapat digunakan untuk menelusuri pola traffic dan alamat IP yang memerlukan pemeriksaan lebih lanjut. Namun, jumlah data, jumlah anomali, dan metrik evaluasi belum dapat ditentukan karena file CSV dan hasil eksekusi model belum tersedia. Hasil penelitian selanjutnya dapat digunakan sebagai dasar pengembangan sistem pemantauan keamanan jaringan berbasis deteksi anomali.

Kata Kunci: *Traffic Jaringan, Deteksi Anomali, Wireshark, Isolation Forest, One-Class SVM, Machine Learning.*

Abstract

The increasing use of Wi-Fi networks has led to higher network traffic density, making continuous monitoring necessary to detect activities that deviate from normal network patterns. This study aims to identify anomalies in Wi-Fi traffic data captured using Wireshark without relying on complete attack labels. Network packets were collected using Wireshark and exported in CSV format for preprocessing. Subsequently, the RequestCount feature was constructed to represent network request activity. Anomaly detection was performed using two unsupervised learning methods, namely Isolation Forest and One-Class SVM. Isolation Forest identifies observations that differ from the majority of the data by isolating them, while One-Class SVM establishes a boundary representing normal data patterns and identifies observations outside this boundary as potential anomalies. The resulting output consists of cleaned traffic data, RequestCount values, anomaly scores, and classifications of Normal or Anomalous. These results can be used to investigate traffic patterns and IP addresses that

require further examination. However, the total number of observations, detected anomalies, and evaluation metrics cannot yet be determined because the CSV dataset and final model execution results are not available. Further implementation and evaluation are expected to support the development of a network security monitoring system based on anomaly detection.

Keywords: *Network Traffic, Anomaly Detection, Wireshark, Isolation Forest, One-Class SVM, Machine Learning.*

✉ Corresponding author :
Email Address : 21410100006@dinamika.ac.id

PENDAHULUAN

Jaringan Wi-Fi banyak digunakan karena memberikan kemudahan akses dan mendukung pertukaran data pada berbagai perangkat secara fleksibel. Peningkatan jumlah perangkat dan aktivitas pengguna menyebabkan traffic jaringan menjadi semakin kompleks sehingga pemantauan diperlukan untuk menjaga kestabilan dan keamanan layanan. Kondisi tersebut dapat menghasilkan pola traffic yang berubah-ubah akibat perbedaan jenis perangkat, aplikasi, intensitas komunikasi, serta karakteristik penggunaan jaringan. Apabila perubahan pola tersebut tidak teridentifikasi dengan baik, gangguan maupun aktivitas yang tidak normal dapat sulit dibedakan dari kondisi traffic normal.

Pemantauan traffic jaringan secara konvensional umumnya bergantung pada aturan, parameter, atau pola yang telah ditentukan sebelumnya. Pendekatan tersebut dapat menghadapi keterbatasan ketika berhadapan dengan karakteristik traffic yang dinamis dan pola anomali yang belum diketahui sebelumnya. Oleh karena itu, deteksi anomali berbasis machine learning menjadi salah satu pendekatan yang banyak dikembangkan karena mampu mempelajari karakteristik data traffic dan mengidentifikasi penyimpangan dari pola normal. Penelitian Ma et al. menunjukkan bahwa metode machine learning dapat digunakan untuk mendeteksi anomali pada traffic jaringan melalui kombinasi Support Vector Machine dan metode clustering (Ma et al., 2021).

Pada lingkungan jaringan nirkabel, kebutuhan terhadap deteksi anomali menjadi semakin penting karena karakteristik jaringan yang dinamis dan adanya berbagai potensi gangguan maupun serangan. Penelitian (Cerdà-Alabern et al., 2023) menunjukkan bahwa pendekatan unsupervised machine learning dapat digunakan untuk mendeteksi kejadian abnormal pada jaringan nirkabel dengan memanfaatkan karakteristik traffic dan informasi kondisi jaringan. Hasil penelitian tersebut menunjukkan bahwa metode deteksi anomali dapat membantu mengidentifikasi gangguan jaringan tanpa harus selalu bergantung pada klasifikasi serangan yang telah ditentukan sebelumnya. Selain itu, karakteristik traffic jaringan yang memiliki variasi pada skala waktu juga menjadi tantangan dalam proses deteksi anomali. Duan et al. mengembangkan metode deteksi anomali traffic jaringan berbasis karakteristik multi-scale dengan memanfaatkan sliding window, transformasi wavelet, dan autoencoder untuk memperoleh informasi pola traffic pada berbagai skala. Pendekatan tersebut menunjukkan bahwa pemodelan karakteristik traffic secara lebih

mendalam dapat membantu meningkatkan kemampuan dalam membedakan traffic normal dan anomali (Duan et al., 2023).

Dengan demikian, analisis traffic Wi-Fi tidak hanya berperan dalam mengetahui tingkat penggunaan jaringan, tetapi juga dapat menjadi bagian penting dalam proses identifikasi kondisi yang menyimpang. Pemanfaatan data hasil tangkapan traffic, seperti data yang diperoleh melalui Wireshark, memberikan peluang untuk mengekstraksi berbagai karakteristik komunikasi jaringan dan menganalisis pola normal maupun abnormal. Pendekatan berbasis deteksi anomali dapat menjadi alternatif ketika data traffic tidak memiliki label serangan yang lengkap, karena proses identifikasi dapat difokuskan pada penyimpangan karakteristik traffic dari pola normal. Perkembangan penelitian dalam bidang network anomaly detection juga menunjukkan semakin luasnya penggunaan metode machine learning untuk meningkatkan kemampuan monitoring dan keamanan jaringan (Liu et al., 2024).

Isu utama pada traffic jaringan adalah munculnya aktivitas yang berbeda dari pola normal, seperti lonjakan permintaan, aktivitas scanning, komunikasi berulang, penggunaan bandwidth yang tidak wajar, atau pola akses yang tidak sesuai dengan karakteristik mayoritas pengguna. Kondisi tersebut dapat dikategorikan sebagai anomali traffic karena menunjukkan adanya penyimpangan terhadap pola komunikasi jaringan yang umumnya terjadi. Anomali tidak selalu menunjukkan serangan siber, tetapi dapat menjadi indikasi awal terjadinya gangguan jaringan, kesalahan konfigurasi perangkat, penggunaan sumber daya secara berlebihan, maupun aktivitas yang berpotensi membahayakan keamanan jaringan. Penelitian mengenai network anomaly detection menunjukkan bahwa perubahan pola traffic dapat dimanfaatkan sebagai indikator untuk mengidentifikasi kondisi abnormal pada jaringan dan mendukung proses monitoring keamanan secara lebih adaptif (Wang et al., 2021).

Pada jaringan Wi-Fi, proses identifikasi anomali memiliki tantangan yang lebih kompleks karena karakteristik traffic dapat berubah mengikuti jumlah perangkat yang terhubung, jenis aplikasi yang digunakan, intensitas pertukaran data, serta kondisi jaringan pada waktu tertentu. Penelitian Cerdà-Alabern et al. menunjukkan bahwa pendekatan unsupervised machine learning dapat digunakan untuk mendeteksi kondisi abnormal pada jaringan nirkabel. Penelitian tersebut menggunakan data dari jaringan nirkabel produksi dan menunjukkan bahwa beberapa metode machine learning mampu mendeteksi kejadian kegagalan jaringan berdasarkan karakteristik traffic dan atribut jaringan lainnya (Cerdà-Alabern et al., 2023).

Variasi karakteristik traffic tersebut menyebabkan penentuan batas antara kondisi normal dan abnormal tidak selalu dapat dilakukan hanya berdasarkan satu parameter jaringan. Oleh karena itu, pendekatan berbasis machine learning menjadi salah satu alternatif untuk mempelajari karakteristik data dan menemukan pola yang menyimpang. Fotiadou et al. menunjukkan bahwa pendekatan deep learning dapat digunakan untuk melakukan deteksi anomali pada traffic jaringan dengan memanfaatkan kemampuan model dalam mempelajari representasi dan pola traffic secara otomatis. Selain itu, (Zuo et al., 2024) mengembangkan metode deteksi anomali berbasis karakteristik multi-scale untuk menangkap informasi traffic pada rentang

waktu yang berbeda, yang menunjukkan bahwa karakteristik temporal traffic dapat menjadi informasi penting dalam proses identifikasi anomali.

Permasalahan lain yang perlu diperhatikan adalah keterbatasan ketersediaan data berlabel. Metode supervised learning umumnya membutuhkan dataset yang telah memiliki label normal dan jenis serangan secara lengkap, sedangkan dalam kondisi nyata tidak seluruh aktivitas abnormal dapat diketahui atau diberi label sebelumnya. Kajian Xu menunjukkan bahwa metode supervised learning sangat bergantung pada ketersediaan dataset berlabel, sedangkan pendekatan unsupervised dan semi-supervised memiliki potensi untuk mengidentifikasi anomali atau serangan baru yang belum memiliki label. Hal tersebut juga menjadi salah satu alasan pendekatan deteksi anomali tanpa ketergantungan penuh terhadap label serangan semakin banyak dikaji dalam penelitian keamanan jaringan (Liu et al., 2024).

Dalam konteks penelitian ini, data hasil packet capture menggunakan Wireshark dapat dimanfaatkan untuk memperoleh karakteristik komunikasi jaringan yang selanjutnya dianalisis untuk menemukan pola normal dan penyimpangan. Informasi seperti jumlah paket, ukuran paket, protokol, alamat sumber dan tujuan, waktu komunikasi, serta karakteristik aliran traffic dapat digunakan sebagai atribut dalam proses analisis. Pendekatan tersebut memungkinkan identifikasi anomali dilakukan berdasarkan karakteristik aktual traffic, bukan hanya berdasarkan kecocokan dengan pola serangan yang telah diketahui sebelumnya. Dengan demikian, deteksi anomali dapat menjadi pendekatan yang relevan untuk membantu proses monitoring traffic Wi-Fi, terutama ketika dataset yang tersedia memiliki keterbatasan label serangan.

Permasalahan yang diangkat dalam artikel ini adalah belum adanya uraian yang sistematis mengenai cara mengolah data Wireshark menjadi informasi yang dapat digunakan untuk mencari anomali traffic Wi-Fi. Data mentah hasil packet capture pada dasarnya terdiri atas sejumlah besar informasi paket dengan karakteristik yang beragam sehingga belum dapat secara langsung digunakan sebagai masukan metode machine learning. Oleh karena itu, data perlu melalui tahapan preprocessing, seperti pembersihan data, penghapusan atribut yang tidak relevan, transformasi data, normalisasi, serta penanganan data yang tidak lengkap. Tahapan preprocessing memiliki peranan penting karena kualitas data masukan dapat memengaruhi kinerja model dalam proses deteksi anomali. (Güney, 2023) menunjukkan bahwa teknik preprocessing dan pemilihan fitur dapat memberikan pengaruh terhadap performa sistem deteksi intrusi berbasis machine learning.

Pemilihan fitur juga menjadi tahap penting karena data traffic jaringan dapat memiliki jumlah atribut yang besar dan tidak seluruh atribut memberikan kontribusi yang sama terhadap proses identifikasi anomali. Penggunaan fitur yang tidak relevan atau memiliki korelasi tinggi dapat meningkatkan kompleksitas model dan berpotensi menurunkan efisiensi proses analisis. Kajian mengenai feature selection dalam sistem deteksi intrusi menunjukkan bahwa pemilihan fitur merupakan bagian penting dari preprocessing karena dapat mengurangi dimensi data sekaligus mempertahankan informasi yang paling relevan untuk proses deteksi (Yu et al., 2024).

Pada artikel ini, Isolation Forest dan One-Class SVM digunakan sebagai dua metode untuk menemukan kandidat anomali pada traffic Wi-Fi. Kedua metode tersebut tidak ditempatkan sebagai objek kompetisi untuk menentukan metode yang

paling unggul, melainkan sebagai pendekatan yang saling melengkapi dalam menghasilkan kandidat anomali berdasarkan karakteristik pola data. Isolation Forest memanfaatkan proses isolasi terhadap observasi yang relatif jarang atau berbeda, sedangkan One-Class SVM membangun batas keputusan terhadap pola data yang dianggap normal. Pendekatan ini memungkinkan proses analisis difokuskan pada identifikasi penyimpangan dari pola traffic normal, bukan semata-mata pada klasifikasi jenis serangan yang telah diketahui sebelumnya. Penggunaan Isolation Forest pada analisis traffic jaringan juga telah diteliti pada skala data besar dan lingkungan jaringan, sedangkan One-Class SVM telah digunakan untuk mendukung deteksi anomali pada lingkungan IoT dan data jaringan yang memiliki karakteristik kompleks (Laskar et al., 2021).

METODOLOGI

Pengumpulan Data

Pengumpulan data dilakukan menggunakan aplikasi Wireshark untuk menangkap traffic jaringan Wi-Fi pada tiga lokasi berbeda di Kota Surabaya yang selanjutnya disajikan sebagai Lokasi A, Lokasi B, dan Lokasi C. Penggunaan kode lokasi dilakukan untuk menjaga kerahasiaan identitas objek penelitian. Pengambilan data dilakukan secara langsung pada jaringan Wi-Fi publik yang digunakan oleh pengguna dalam aktivitas sehari-hari. Ketiga lokasi penelitian merupakan area publik yang menyediakan akses internet nirkabel (Wi-Fi) bagi pengunjung. Traffic yang terekam berasal dari berbagai aktivitas pengguna, seperti penelusuran web (browsing), penggunaan media sosial, layanan pesan instan, streaming video, serta aktivitas komunikasi data lainnya. Variasi aktivitas tersebut menghasilkan pola traffic yang beragam sehingga dapat digunakan sebagai sumber data untuk identifikasi anomali jaringan.

Proses packet capture dilakukan selama kurang lebih tiga jam pada masing-masing lokasi menggunakan Wireshark. Seluruh paket yang melewati jaringan selama periode pengamatan direkam dan disimpan dalam format packet capture, kemudian diekspor ke format Comma Separated Values (CSV) untuk memudahkan proses analisis lebih lanjut. Dataset yang diperoleh memiliki tujuh atribut utama, yaitu No., Time, Source, Destination, Protocol, Length, dan Info. Hasil pengumpulan data menunjukkan bahwa Lokasi A menghasilkan 18.634 record traffic jaringan, Lokasi B menghasilkan 31.863 record traffic jaringan, sedangkan Lokasi C menghasilkan 764.530 record traffic jaringan. Secara keseluruhan diperoleh 815.027 record traffic jaringan yang digunakan sebagai data awal penelitian.

Data hasil capture kemudian diproses lebih lanjut melalui tahap preprocessing dan pembentukan fitur RequestCount. Setelah dilakukan agregasi berdasarkan atribut Source, diperoleh 109 observasi pada Lokasi A, 177 observasi pada Lokasi B, dan 399 observasi pada Lokasi C. Dengan demikian total observasi yang digunakan pada proses identifikasi anomali sebanyak 685 observasi yang selanjutnya dianalisis menggunakan metode Isolation Forest dan One-Class SVM.

Preprocessing dan Pembentukan Fitur

Tahap preprocessing dilakukan untuk menyiapkan data hasil packet capture agar dapat digunakan dalam proses identifikasi anomali. Dataset hasil ekspor

Wireshark dalam format CSV terdiri atas atribut No., Time, Source, Destination, Protocol, Length, dan Info. Sebelum dilakukan analisis, struktur data diperiksa untuk memastikan seluruh atribut dapat dibaca dengan baik oleh sistem. Tahap ini bertujuan untuk menghindari kesalahan pemrosesan yang dapat memengaruhi hasil identifikasi anomali.Pemilihan atribut dilakukan berdasarkan relevansinya terhadap tujuan penelitian, yaitu mendeteksi aktivitas jaringan yang menyimpang dari pola normal. Dari seluruh atribut yang tersedia, atribut Source dipilih sebagai atribut utama karena menunjukkan alamat sumber yang menghasilkan traffic jaringan. Dalam penelitian deteksi anomali jaringan, frekuensi kemunculan suatu alamat sumber dapat digunakan untuk menggambarkan tingkat aktivitas sebuah host pada jaringan. Host yang menghasilkan traffic jauh lebih tinggi atau jauh lebih rendah dibanding mayoritas host lainnya berpotensi menunjukkan perilaku yang tidak normal sehingga perlu dilakukan pemeriksaan lebih lanjut.

Tabel 1. Atribut Yang Digunakan Dalam Penelitian

Atribut	Keterangan	Digunakan
Source	Alamat sumber paket	Ya
Destination	Alamat tujuan paket	Tidak
Protocol	Jenis protokol komunikasi	Tidak
Length	Ukuran paket	Tidak
Time	Waktu paket ditangkap	Tidak
RequestCount	Jumlah kemunculan Source	Ya

Berdasarkan Tabel 1, atribut *Source* dipilih sebagai atribut utama dalam penelitian ini. Atribut tersebut digunakan untuk menghitung frekuensi kemunculan setiap alamat sumber pada data traffic jaringan. Hasil perhitungan frekuensi kemudian direpresentasikan dalam bentuk fitur *Request Count* yang digunakan sebagai masukan bagi model *Isolation Forest* dan *One-Class* SVM untuk mengidentifikasi anomali pada traffic jaringan Wi-Fi. Setelah atribut *Source* ditentukan, dilakukan proses feature engineering untuk membentuk fitur RequestCount. Fitur *Request Count* merepresentasikan jumlah kemunculan setiap alamat Source pada dataset hasil capture. Nilai ini digunakan sebagai indikator tingkat aktivitas suatu host selama periode pengamatan. Semakin besar nilai RequestCount, semakin tinggi jumlah aktivitas yang dilakukan oleh alamat sumber tersebut pada jaringan. Perhitungan RequestCount dilakukan dengan menghitung frekuensi kemunculan setiap alamat Source pada seluruh data traffic. Secara matematis, nilai RequestCount dapat dirumuskan sebagai berikut:

$$\text{RequestCount}_i = \Sigma\, I(\text{Source}_j = \text{Source}_i)$$

dengan:

1. RequestCount_i adalah jumlah aktivitas dari alamat sumber ke-i.
2. I merupakan fungsi indikator yang bernilai 1 apabila Source_j sama dengan Source_i dan bernilai 0 apabila berbeda.
3. j menunjukkan setiap paket yang terdapat dalam dataset.

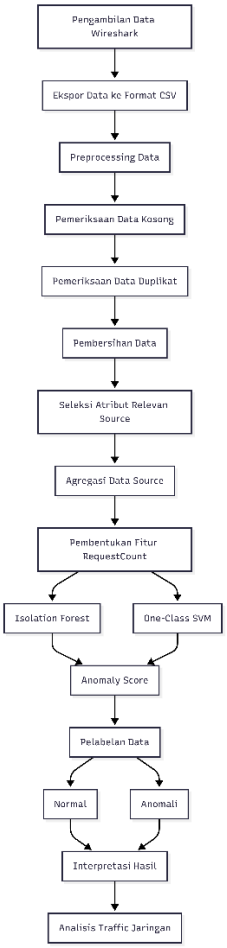
Sebagai contoh, apabila data traffic yang diperoleh terdiri atas lima paket dengan alamat sumber seperti pada Tabel 1, maka nilai RequestCount dapat dihitung berdasarkan jumlah kemunculan masing-masing alamat sumber.

Tabel 2. Contoh pembentukan fitur RequestCount

Source	Jumlah Kemunculan
192.168.1.129	3
92.223.76.254	2

Berdasarkan Tabel 2, alamat 192.168.1.129 muncul sebanyak tiga kali sehingga memiliki nilai RequestCount sebesar 3. Sementara itu, alamat 92.223.76.254 muncul sebanyak dua kali sehingga memiliki nilai RequestCount sebesar 2. Hasil agregasi tersebut menghasilkan dua atribut utama, yaitu Source dan RequestCount. Fitur RequestCount selanjutnya digunakan sebagai masukan bagi metode Isolation Forest dan One-Class SVM untuk mengidentifikasi observasi yang memiliki pola aktivitas berbeda dari mayoritas traffic jaringan. Fitur RequestCount selanjutnya digunakan sebagai masukan bagi metode Isolation Forest dan One-Class SVM untuk mengidentifikasi observasi yang memiliki pola aktivitas berbeda dari mayoritas traffic jaringan. Proses pembentukan fitur RequestCount pada penelitian ini dapat dilihat pada Gambar 1.

Gambar 1. Proses Pembentukan Fitur RequestCount



Berdasarkan Gambar 1, fitur RequestCount dibentuk melalui proses agregasi alamat Source pada data hasil capture Wireshark. Frekuensi kemunculan setiap alamat Source dihitung dan digunakan sebagai representasi tingkat aktivitas host pada jaringan.

Identifikasi Menggunakan Isolation Forest

Isolation Forest merupakan metode deteksi anomali berbasis unsupervised learning yang bekerja dengan cara mengisolasi observasi yang berbeda dari mayoritas data. Metode ini dikembangkan berdasarkan asumsi bahwa data anomali memiliki karakteristik yang lebih mudah dipisahkan dibandingkan data normal karena jumlahnya relatif sedikit dan memiliki nilai yang berbeda secara signifikan dari sebagian besar observasi lainnya. Pada penelitian ini, Isolation Forest digunakan untuk mengidentifikasi anomali berdasarkan fitur RequestCount yang telah dibentuk pada tahap sebelumnya. Nilai RequestCount digunakan sebagai representasi tingkat aktivitas setiap alamat Source dalam jaringan. Host yang memiliki jumlah aktivitas sangat tinggi atau sangat rendah dibandingkan mayoritas host lainnya berpotensi teridentifikasi sebagai anomali.

Proses identifikasi dilakukan dengan membangun sejumlah pohon isolasi (isolation trees) secara acak. Setiap observasi dipisahkan melalui serangkaian proses pembagian data hingga mencapai kondisi terisolasi. Observasi yang membutuhkan jalur pemisahan lebih pendek cenderung dikategorikan sebagai anomali karena lebih mudah dipisahkan dari kumpulan data lainnya. Secara matematis, tingkat keanomalian suatu observasi pada Isolation Forest dapat direpresentasikan menggunakan anomaly score sebagai berikut:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$$

dengan:

1. E(h(x)) merupakan rata-rata panjang jalur (path length) yang dibutuhkan untuk mengisolasi observasi x pada seluruh pohon isolasi.
2. c(n) merupakan faktor normalisasi berdasarkan jumlah observasi n.
3. x merupakan observasi yang dianalisis.

Nilai anomaly score digunakan untuk mengukur tingkat penyimpangan suatu observasi terhadap mayoritas data. Semakin pendek jalur isolasi yang dibutuhkan untuk memisahkan suatu observasi dari kumpulan data lainnya, semakin besar kemungkinan observasi tersebut merupakan anomali. Pada implementasi penelitian ini, observasi yang menghasilkan nilai score lebih rendah cenderung memiliki tingkat penyimpangan yang lebih tinggi dibandingkan observasi lainnya. Pada penelitian ini, nilai RequestCount yang jauh berbeda dari mayoritas observasi berpotensi menghasilkan anomaly score yang menunjukkan penyimpangan sehingga dapat diklasifikasikan sebagai Anomali oleh model Isolation Forest. Sebaliknya, observasi dengan nilai RequestCount yang masih mengikuti pola mayoritas data cenderung memperoleh score yang menunjukkan kondisi normal.

Pada implementasi penelitian ini, model Isolation Forest dibangun menggunakan parameter contamination sebesar 0,2 dan random_state sebesar 42. Parameter contamination digunakan untuk memperkirakan proporsi data yang berpotensi merupakan anomali dalam dataset, sedangkan random_state digunakan untuk menjaga konsistensi hasil pengujian. Hasil keluaran model berupa anomaly score dan nilai prediksi. Nilai prediksi 1 menunjukkan bahwa observasi termasuk dalam kategori Normal, sedangkan nilai prediksi -1 menunjukkan bahwa observasi termasuk dalam kategori Anomali. Pada penelitian ini, nilai prediksi tersebut

dikonversi menjadi label Normal dan Anomali untuk memudahkan proses interpretasi hasil. Observasi yang memperoleh label Anomali menunjukkan pola aktivitas yang berbeda dari mayoritas traffic jaringan sehingga memerlukan pemeriksaan lebih lanjut oleh administrator jaringan.

Secara umum, proses identifikasi menggunakan Isolation Forest pada penelitian ini dapat diringkas sebagai berikut:

1. Data hasil agregasi RequestCount digunakan sebagai masukan model.
2. Model membangun pohon isolasi secara acak untuk setiap observasi.
3. Model menghitung anomaly score berdasarkan tingkat kemudahan suatu observasi untuk diisolasi.
4. Observasi diberi label Normal atau Anomali berdasarkan hasil perhitungan model.
5. Hasil akhir berupa alamat Source, nilai RequestCount, anomaly score, dan status identifikasi.

HASIL DAN PEMBAHASAN

Hasil proses dari dokumen awal menunjukkan bahwa data Wireshark telah diarahkan untuk diolah menjadi fitur RequestCount dan dianalisis menggunakan Isolation Forest serta One-Class SVM. Keluaran utama yang dibahas adalah nilai RequestCount, score, dan label status Normal atau Anomali. Data yang dihasilkan dapat dibagi menjadi tiga kelompok, yaitu data input, data hasil preprocessing, dan data output model. Data input berasal dari packet capture Wireshark, data preprocessing berupa observasi yang telah dibersihkan dan diringkas, sedangkan data output model berupa score dan label yang digunakan untuk menandai kandidat anomali. Pada dokumen awal, jumlah anomali dari Isolation Forest masih ditulis sebagai X dan jumlah anomali dari One-Class SVM masih ditulis sebagai Y. Hal ini menunjukkan bahwa hasil numerik belum final dan perlu diperbarui berdasarkan eksekusi model pada file CSV asli agar uraian hasil memenuhi standar artikel jurnal.

Hasil Identifikasi Anomali

Penelitian ini menggunakan data traffic jaringan Wi-Fi yang diperoleh dari tiga lokasi pengamatan berbeda di Kota Surabaya. Data hasil packet capture terlebih dahulu diproses melalui tahap preprocessing dan pembentukan fitur RequestCount sebelum dianalisis menggunakan metode Isolation Forest dan One-Class SVM. Hasil preprocessing menghasilkan 109 observasi pada Lokasi A, 177 observasi pada Lokasi B, dan 399 observasi pada Lokasi C. Dengan demikian, total observasi yang digunakan dalam proses identifikasi anomali sebanyak 685 observasi.

Tabel 3. Jumlah observasi hasil preprocessing

Lokasi Pengamatan	Jumlah Observasi
Lokasi A	109
Lokasi B	177
Lokasi C	399
Total	685

Selanjutnya setiap observasi dianalisis menggunakan Isolation Forest dan One-Class SVM. Hasil analisis menghasilkan nilai RequestCount, score, serta label Normal atau Anomali untuk setiap alamat Source yang diamati. Pada Lokasi A, nilai

RequestCount tertinggi ditemukan pada alamat sumber 192.168.1.239 dengan nilai RequestCount sebesar 7.336. Pada Lokasi B, nilai RequestCount tertinggi ditemukan pada alamat sumber 10.214.79.161 dengan nilai RequestCount sebesar 13.736. Sementara itu, pada Lokasi C nilai RequestCount tertinggi ditemukan pada alamat sumber 192.168.1.129 dengan nilai RequestCount sebesar 285.705. Hasil tersebut menunjukkan bahwa setiap lokasi memiliki karakteristik traffic yang berbeda. Lokasi C memiliki tingkat aktivitas jaringan yang jauh lebih tinggi dibandingkan dua lokasi lainnya, yang terlihat dari besarnya nilai RequestCount pada beberapa alamat sumber utama. Perbedaan karakteristik traffic ini menunjukkan pentingnya penggunaan metode deteksi anomali yang mampu menyesuaikan diri dengan distribusi data pada masing-masing lokasi.

Secara umum, kedua metode berhasil menghasilkan label Normal dan Anomali pada setiap observasi. Label Anomali menunjukkan bahwa suatu observasi memiliki pola aktivitas yang berbeda dari mayoritas traffic pada lokasi yang sama, sedangkan label Normal menunjukkan bahwa aktivitas tersebut masih berada dalam pola yang dianggap wajar oleh model.3.2 Data Turunan dan Hubungan dengan Permasalahan

Hasil Identifikasi Menggunakan Isolation Forest

Proses identifikasi anomali menggunakan Isolation Forest dilakukan terhadap seluruh observasi hasil preprocessing yang telah direpresentasikan dalam bentuk fitur RequestCount. Metode ini menghasilkan dua keluaran utama, yaitu anomaly score dan label identifikasi yang terdiri atas kategori Normal dan Anomali. Pada Lokasi A, observasi dengan RequestCount tertinggi berasal dari alamat sumber 192.168.1.239 dengan nilai RequestCount sebesar 7.336 dan memperoleh label Anomali. Selain itu, beberapa alamat sumber lain seperti 172.64.155.209, 31.13.95.60, dan 146.75.46.172 juga memperoleh label Anomali karena memiliki tingkat aktivitas yang berbeda dari mayoritas observasi pada lokasi tersebut.

Pada Lokasi B, nilai RequestCount tertinggi ditemukan pada alamat sumber 10.214.79.161 dengan nilai RequestCount sebesar 13.736. Observasi tersebut memperoleh label Anomali dengan anomaly score negatif, yang menunjukkan adanya penyimpangan terhadap pola mayoritas traffic jaringan. Beberapa observasi lain dengan nilai RequestCount tinggi juga memperoleh label Anomali karena memiliki karakteristik yang berbeda dibandingkan sebagian besar data. Pada Lokasi C, nilai RequestCount tertinggi ditemukan pada alamat sumber 192.168.1.129 dengan nilai RequestCount sebesar 285.705. Selain itu, alamat sumber 92.223.76.254, 2404:c0:b602:3818:2910:69ea:7b79:f0f1, dan 146.75.46.172 juga memperoleh label Anomali. Nilai RequestCount yang jauh lebih besar dibandingkan mayoritas observasi menyebabkan alamat-alamat tersebut lebih mudah diisolasi oleh model sehingga dikategorikan sebagai kandidat anomali.

Hasil identifikasi menunjukkan bahwa Isolation Forest mampu menemukan observasi dengan tingkat aktivitas yang berbeda secara signifikan dari mayoritas traffic jaringan. Observasi yang memperoleh label Anomali tidak secara langsung dianggap sebagai serangan, namun digunakan sebagai indikator awal untuk menentukan traffic yang perlu diperiksa lebih lanjut oleh administrator jaringan.

Secara umum, observasi dengan nilai RequestCount sangat tinggi cenderung memperoleh anomaly score yang lebih rendah dan lebih sering dikategorikan sebagai Anomali. Hal ini sesuai dengan prinsip kerja Isolation Forest yang menganggap observasi dengan karakteristik berbeda dari mayoritas data lebih mudah dipisahkan dalam struktur pohon isolasi yang dibangun oleh model.

Hasil Identifikasi Menggunakan One-Class SVM

Selain menggunakan Isolation Forest, penelitian ini juga menerapkan metode One-Class SVM untuk mengidentifikasi observasi yang menyimpang dari pola mayoritas traffic jaringan. Metode ini menghasilkan keluaran berupa decision score dan label identifikasi yang terdiri atas kategori Normal dan Anomali. Pada Lokasi A, alamat sumber 192.168.1.239 yang memiliki nilai RequestCount sebesar 7.336 memperoleh label Normal. Namun demikian, beberapa alamat sumber lain seperti 172.64.155.209, 31.13.95.60, dan 146.75.46.172 memperoleh label Anomali karena berada di luar batas keputusan yang dibentuk oleh model. Hasil ini menunjukkan bahwa nilai RequestCount yang tinggi tidak selalu dianggap sebagai anomali apabila masih mengikuti pola distribusi data yang dipelajari oleh One-Class SVM.

Pada Lokasi B, alamat sumber 10.214.79.161 dengan RequestCount sebesar 13.736 memperoleh label Anomali. Sementara itu, beberapa observasi lain dengan nilai RequestCount yang relatif tinggi masih memperoleh label Normal. Kondisi tersebut menunjukkan bahwa proses identifikasi tidak hanya dipengaruhi oleh besarnya nilai RequestCount, tetapi juga oleh posisi observasi terhadap batas keputusan yang dibentuk model. Pada Lokasi C, alamat sumber dengan RequestCount tertinggi sebesar 285.705 memperoleh label Normal. Hasil yang sama juga ditemukan pada beberapa observasi lain dengan nilai RequestCount yang besar. Temuan ini menunjukkan bahwa One-Class SVM menganggap observasi tersebut masih berada dalam pola mayoritas data sehingga tidak dikategorikan sebagai anomali.

Secara umum, One-Class SVM menghasilkan interpretasi yang berbeda dibandingkan Isolation Forest pada beberapa observasi. Perbedaan tersebut muncul karena kedua metode menggunakan pendekatan yang berbeda dalam mendeteksi anomali. Isolation Forest mengidentifikasi anomali berdasarkan kemudahan suatu observasi untuk diisolasi, sedangkan One-Class SVM mengidentifikasi anomali berdasarkan posisi observasi terhadap batas keputusan yang dibentuk dari pola data normal. Hasil identifikasi menggunakan One-Class SVM memberikan tambahan informasi bagi proses analisis traffic jaringan. Observasi yang memperoleh label Anomali dapat digunakan sebagai kandidat traffic yang perlu diperiksa lebih lanjut, sedangkan observasi yang memperoleh label Normal menunjukkan bahwa aktivitas tersebut masih dianggap mengikuti pola mayoritas jaringan berdasarkan model yang digunakan.

Dampak Hasil Analisis

Dampak praktis dari hasil analisis adalah administrator jaringan memperoleh daftar traffic yang perlu diperiksa lebih dahulu berdasarkan label Anomali, RequestCount, dan anomaly score.[1][7] Pemeriksaan lanjutan dapat diarahkan pada alamat IP, waktu kejadian, protokol, dan pola komunikasi untuk memastikan apakah anomali tersebut berkaitan dengan gangguan atau aktivitas berbahaya.Dampak akademik artikel ini adalah tersedianya alur penulisan dan analisis yang menjelaskan cara mengubah data Wireshark menjadi data analitik untuk deteksi anomali berbasis machine learning.[3][9] Alur ini dapat digunakan sebagai dasar pengembangan penelitian berikutnya dengan menambahkan dataset aktual, validasi label, dan pengujian metrik evaluasi. Keterbatasan utama artikel ini adalah belum tersedianya angka final seperti total record, jumlah anomali, precision, recall, dan F1-score.[9] Oleh karena itu, artikel perlu dilengkapi dengan file CSV asli atau output model agar bagian hasil dan pembahasan dapat diisi dengan angka yang dapat diverifikasi.

SIMPULAN

Artikel ini menyajikan format penulisan jurnal untuk identifikasi anomali traffic data jaringan Wi-Fi menggunakan Wireshark, RequestCount, Isolation Forest, dan One-Class SVM. Proses identifikasi dilakukan melalui pengumpulan data, ekspor CSV, preprocessing, pembentukan fitur, penerapan model, pembentukan anomaly score, dan pemberian label Normal atau Anomali. Hasil yang diharapkan berupa data utama dan data turunan yang dapat digunakan untuk menelusuri traffic mencurigakan berdasarkan RequestCount, score, dan status anomali. Artikel ini tidak bertujuan membandingkan kedua metode, melainkan menggunakan Isolation Forest dan One-Class SVM sebagai dua cara untuk mencari kandidat anomali yang membutuhkan pemeriksaan lanjutan. Untuk menyempurnakan artikel, penulis perlu memasukkan data CSV atau output model agar jumlah data, jumlah anomali, persentase anomali, precision, recall, dan F1-score dapat ditulis secara faktual.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada pengelola jurnal Mirai atas kesempatannya dalam melakukan proses artikel ini. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang telah memberikan dukungan, masukan, dan bantuan dalam proses pengumpulan data, pengolahan data, serta penyusunan artikel ini. Dukungan dan kontribusi tersebut sangat membantu dalam penyelesaian penelitian dan penyusunan artikel ini.

Referensi

Cerdà-Alabern, L., Iuhasz, G., & Gemmi, G. (2023). Anomaly detection for fault detection in wireless community networks using machine learning. *Computer Communications*, *202*. https://doi.org/10.1016/j.comcom.2023.02.019

Duan, X., Fu, Y., & Wang, K. (2023). Network traffic anomaly detection method based on multi-scale residual classifier. *Computer Communications*, *198*. https://doi.org/10.1016/j.comcom.2022.10.024

Güney, H. (2023). Preprocessing Impact Analysis for Machine Learning-Based Network Intrusion Detection. *Sakarya University Journal of Computer and Information Sciences*, *6*(1). https://doi.org/10.35377/saucis...1223054

Laskar, M. T. R., Huang, J. X., Smetana, V., Stewart, C., Pouw, K., An, A., Chan, S., & Liu, L. (2021). Extending isolation forest for anomaly detection in big data via K-means. *ACM Transactions on Cyber-Physical Systems*, *5*(4). https://doi.org/10.1145/3460976

Liu, R., Shi, J., Chen, X., & Lu, C. (2024). Network anomaly detection and security defense technology based on machine learning: A review. *Computers and Electrical Engineering*, *119*. https://doi.org/10.1016/j.compeleceng.2024.109581

Ma, Q., Sun, C., & Cui, B. (2021). A Novel Model for Anomaly Detection in Network Traffic Based on Support Vector Machine and Clustering. *Security and Communication Networks*, *2021*. https://doi.org/10.1155/2021/2170788

Wang, S., Balarezo, J. F., Kandeepan, S., Al-Hourani, A., Chavez, K. G., & Rubinstein, B. (2021). Machine learning in network anomaly detection: A survey. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3126834

Yu, X., Huang, Y., Zhang, Y., Song, M., & Jia, Z. (2024). Network Intrusion Traffic Detection Based on Feature Extraction. *Computers, Materials and Continua*, *78*(1). https://doi.org/10.32604/cmc.2023.044999

Zuo, F., Zhang, D., Li, L., He, Q., & Deng, J. (2024). GSOOA-1DDRSN: Network traffic anomaly detection based on deep residual shrinkage networks. *Heliyon*, *10*(11). https://doi.org/10.1016/j.heliyon.2024.e32087