Series: Engineering and Science Vol. 10 No. 1 (2024) E-ISSN: 2621-9794, P-ISSN: 2477-2097

PENGEMBANGAN MODEL *DECISION TREE* UNTUK SERANGAN *DISTRIBUTED DENIAL OF SERVICE*

Agus Tedyyana[1)], Afis Julianto[2)], Dedi Hermawan[3)], M Afridon[4)], Faisal Riza[5)]

[1,2,3,5]Jurusan Teknik Informatika, Politeknik Negeri Bengkalis, [4]Jurusan Teknik Elektro, Politeknik Negeri Bengkalis
E-mail : agustedyyana@polbeng.ac.id

Abstract

In the growing digital age, distributed denial of service (DDoS) attacks have become one of the most pressing and destructive cyber security threats. To address this, the research developed and implemented the Decision Tree model to detect DDoS attacks effectively. An intrusion detection system integrates the model, utilizing machine learning technology to analyze TCP data flows in real-time, with the aim of enhancing network detection capabilities and bolstering security measures against DDoS attacks. We built the Decision Tree model using the NF-UQ-NIDS dataset, which includes network traffic data representative of both DDoS attacks and normal traffic. Early data analysis using Wireshark provides additional insight into attack patterns, which helps with model calibration and validation. The developed system effectively identified attacks and sent real-time notifications via Telegram, facilitating prompt action from the security team. The results of this study show that the integration of machine learning into network security systems offers a significant improvement in the speed and accuracy of attack detection, showing enormous potential for further applications in a dynamic and diverse environment. Recommendations for further research include developing hybrid algorithms, implementing automated responses, and expanding notification platforms to strengthen the cyber security architecture against DDoS attacks and similar threats.

Keywords: *Cyber security, DDoS, Decision Tree, Machine Learning, Intrusion Detection,*

PENDAHULUAN

Di tengah kemajuan teknologi informasi dan komunikasi yang semakin pesat, Indonesia menghadapi tantangan besar dalam mengamankan infrastruktur digitalnya dari serangan siber (Budiman, 2022), khususnya serangan *Distributed Denial of Service* (DDoS) (Tsobdjou et al., 2022). Serangan DDoS, yang secara strategis membanjiri server dengan permintaan berlebih sehingga layanan menjadi tidak tersedia, telah menjadi senjata umum bagi pelaku kejahatan siber untuk mengganggu operasional bisnis serta layanan publik(Black & Kim, 2022). Gangguan ini tidak hanya berdampak pada kelancaran operasional tetapi juga menimbulkan kerugian finansial yang signifikan dan mengurangi kepercayaan masyarakat terhadap platform digital. Sebagai contoh, beberapa insiden besar di Indonesia menunjukkan kerentanan sistem-sistem penting negara terhadap serangan DDoS. Serangan terhadap institusi keuangan besar beberapa

tahun yang lalu, yang membuat layanan perbankan online tidak tersedia selama berjam-jam, adalah salah satu kasus yang menyoroti perlunya sistem deteksi dan respons yang lebih baik (P et al., 2021). Insiden tersebut tidak hanya menghambat transaksi finansial tetapi juga menyebabkan penurunan kepercayaan pelanggan yang berkepanjangan. Kejadian ini memperjelas bahwa serangan DDoS dapat mempengaruhi lebih dari sekedar layanan teknis; mereka juga memiliki potensi untuk mengganggu kestabilan ekonomi dan keamanan social (Rao & Subbarao, 2023). Menghadapi realitas ini, kebutuhan akan sistem keamanan siber yang canggih dan responsif sangat mendesak. Pendekatan tradisional yang banyak mengandalkan firewall dan software antivirus terbukti tidak cukup menghadapi taktik yang semakin kompleks dari serangan DDoS. Dalam konteks ini, teknologi pembelajaran mesin menawarkan solusi yang potensial, khususnya melalui penggunaan model *Decision Tree* (Hernández et al., 2022). Model ini, yang menganalisis pola lalu lintas data untuk mengidentifikasi perilaku tidak normal, diharapkan dapat membantu dalam mendeteksi dan merespons serangan dengan lebih cepat dan akurat.

Penelitian ini bertujuan untuk mengembangkan model *Decision Tree* yang dapat secara efisien mendeteksi serangan DDoS dalam lalu lintas jaringan. Dengan memanfaatkan dataset NF-UQ-NIDS (Ma et al., 2023), yang menyediakan sampel realistis dari lalu lintas jaringan normal dan lalu lintas selama serangan DDoS, model ini diharapkan tidak hanya dapat mengidentifikasi serangan tetapi juga mempelajari berbagai teknik serangan yang mungkin berkembang. Selanjutnya, model yang dikembangkan akan diintegrasikan ke dalam sistem deteksi intrusi yang ada untuk meningkatkan kemampuan mereka dalam memonitor, mendeteksi, dan merespons serangan secara real-time. Implementasi ini termasuk pemberitahuan segera melalui sistem komunikasi seperti Telegram (Tedyyana et al., 2024), yang akan memberikan waktu respon yang cepat bagi administrator jaringan untuk mengambil tindakan pencegahan. Penelitian ini tidak hanya relevan dari segi teknologi tetapi juga dalam konteks sosio-ekonomi, mengingat dampak luas dari serangan siber pada keamanan nasional dan ekonomi digital Indonesia. Dengan mengevaluasi kinerja model dalam kondisi simulasi dan nyata, serta mengintegrasikannya dalam infrastruktur yang ada,

diharapkan dapat memberikan kontribusi yang signifikan dalam memperkuat pertahanan siber Indonesia.

METODE PENELITIAN

Penelitian ini berfokus pada pengembangan dan evaluasi model *Decision Tree* untuk deteksi serangan DDoS. Pendekatan kuantitatif dipilih karena memungkinkan pengukuran objektif dari kinerja model berdasarkan data lalu lintas jaringan yang dikuantifikasi. Pendekatan ini melibatkan analisis statistik yang mendalam untuk menguji hipotesis bahwa model *Decision Tree* dapat secara efektif mengidentifikasi dan memprediksi serangan DDoS dari data lalu lintas yang kompleks.

Populasi yang digunakan adalah lalu lintas jaringan yang mengalami serangan DDoS. Data untuk populasi ini diambil dari dataset NF-UQ-NIDS, yang menyediakan data lalu lintas jaringan secara real-time dan telah dianotasi untuk membedakan antara lalu lintas normal dan serangan DDoS. Sampel penelitian dipilih melalui metode sampling acak stratifikasi untuk memastikan bahwa setiap jenis serangan DDoS yang dicakup dalam dataset diwakili secara proporsional dalam data yang digunakan untuk mengembangkan model. Variabel-variabel yang akan dianalisis dalam penelitian ini meliputi:

1. Lalu Lintas Jaringan, Ini termasuk semua data yang diperoleh dari dataset NF-UQ-NIDS. Data ini terdiri dari paket-paket yang dikirim dan diterima melalui jaringan selama periode tertentu, termasuk header dan payload yang memungkinkan identifikasi sumber, tujuan, dan jenis lalu lintas.
2. Serangan DdoS, Ini adalah kategori dari lalu lintas jaringan yang menunjukkan apakah lalu lintas tersebut merupakan bagian dari serangan DDoS atau tidak. Kategori ini ditentukan berdasarkan anotasi dalam dataset NF-UQ-NIDS.
3. Efektivitas Deteksi, Diukur melalui metrik seperti akurasi, presisi, recall, dan F1-score, yang menunjukkan kemampuan model dalam mengklasifikasikan lalu lintas sebagai bagian dari serangan DDoS atau lalu lintas normal.

Data dikumpulkan langsung dari dataset NF-UQ-NIDS yang terbuka untuk penelitian keamanan siber. Dataset ini mencakup informasi detail tentang lalu lintas jaringan di bawah kondisi normal serta selama serangan DDoS. Data ini sudah tersedia

dalam format yang bisa langsung digunakan untuk analisis machine learning, termasuk atribut-atribut yang relevan seperti IP sumber, IP tujuan, protokol, panjang paket, dan timestamp. Setelah data terkumpul, dilakukan pra-pemrosesan untuk menghilangkan noise dan menghandle missing data. Proses ini mencakup normalisasi data, pengkodean variabel kategorikal, dan pemisahan dataset menjadi set pelatihan dan pengujian untuk menghindari overfitting dan untuk validasi model. Analisis data dalam penelitian ini melibatkan beberapa tahap:

1. Pra-pemrosesan Data, Ini melibatkan penanganan data yang hilang, penghapusan duplikat, dan transformasi data (seperti normalisasi) untuk membuat data siap untuk analisis (Tedyyana et al., 2024).
2. Pembangunan Model *Decision Tree*, Model ini dikonfigurasi dengan menggunakan set pelatihan yang telah diproses. Selama fase ini, dilakukan eksperimen dengan berbagai parameter model untuk menemukan konfigurasi terbaik yang menghasilkan prediksi paling akurat (Kowal, 2022).
3. Validasi Model, Model yang telah dibangun diuji menggunakan set pengujian untuk menilai kinerjanya dalam kondisi yang belum pernah dilihat sebelumnya. Kinerja model diukur menggunakan metrik seperti akurasi, presisi, recall, dan F1-score (Yacouby & Axman, 2020).
4. Optimasi Model, Berdasarkan hasil dari validasi, model mungkin perlu di-tune lebih lanjut untuk meningkatkan kinerjanya. Ini bisa melibatkan penyesuaian parameter, seleksi fitur, atau teknik resampling untuk mengatasi ketidakseimbangan kelas dalam data (Zhang et al., 2022).
5. Implementasi dan Uji Coba di Lingkungan Nyata, Setelah model tervalidasi dan dioptimasi, langkah terakhir adalah implementasinya dalam sistem deteksi intrusi yang ada. Model diintegrasikan dan diuji dalam lingkungan operasional untuk menilai efektivitasnya dalam kondisi nyata dan melakukan penyesuaian berdasarkan feedback yang diterima.

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan dataset yang secara khusus menggambarkan berbagai jenis serangan siber, termasuk serangan DDoS, serangan injection, dan XSS. Dari data

yang diamati, terdapat sejumlah besar entri yang mencatat jenis-jenis serangan tersebut, serta lalu lintas jaringan yang bersifat benign (normal). Ini menunjukkan variasi yang luas dalam jenis serangan dan frekuensi mereka, memberikan basis data yang kaya untuk mengembangkan dan menguji model *Decision Tree* dalam mendeteksi serangan siber.

```
1 # check the number of values for labels
2 network_data['Attack'].value_counts()
```

```
Attack
Benign            25165295
DDoS              21748351
DoS               17875585
scanning           3781419
Reconnaissance     2633778
xss                2455020
password           1153323
injection           684897
Bot                 143097
Brute Force         123982
Infilteration       116361
Exploits             31551
Fuzzers              22310
Backdoor             18978
Generic              16560
mitm                  7723
ransomware            3425
```

Gambar 1. Analisis awal dataset

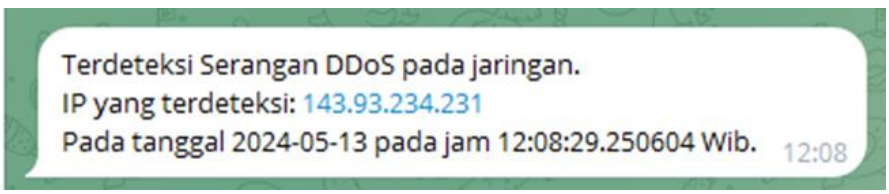
Berdasarkan analisis awal dari dataset pada gambar 1, serangan DDoS mendominasi dengan jumlah mencapai lebih dari 21 juta entri, yang menunjukkan frekuensi serangan yang hampir setara dengan lalu lintas normal. Ini menandakan bahwa DDoS adalah ancaman yang sangat sering terjadi dan memerlukan perhatian khusus dalam pengembangan strategi deteksi serangan. Serangan jenis lain seperti injection dan XSS juga tercatat dalam jumlah signifikan, tetapi dalam skala yang lebih kecil jika dibandingkan dengan DDoS dan lalu lintas normal.

```
Accuracy: 0.9992
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00    137735
           1       1.00      1.00      1.00    114879

    accuracy                           1.00    252614
   macro avg       1.00      1.00      1.00    252614
weighted avg       1.00      1.00      1.00    252614
```

Gambar 2. Klasifikasi Model *Decision Tree*

Dari gambar 2. Klasifikasi model *Decision Tree* menunjukan Model ini mencapai tingkat akurasi yang sangat tinggi yaitu 99.92%. Ini mengindikasikan bahwa model hampir sempurna dalam mengidentifikasi serangan DDoS dan lalu lintas normal, memvalidasi keefektivitasan model dalam kondisi pengujian yang diatur selama fase pengembangan

Dalam evaluasi model *Decision Tree* yang dikembangkan, metrik seperti akurasi, presisi, recall, dan F1-score digunakan untuk menilai kinerjanya. Dari pengujian awal, model menunjukkan kecenderungan yang baik dalam mengklasifikasikan serangan DDoS dibandingkan dengan jenis serangan lainnya dan lalu lintas benign, menunjukkan bahwa model dapat dengan efektif membedakan antara serangan dan lalu lintas normal. Hal ini kritikal dalam pengaturan operasional dimana deteksi dini dan akurat serangan DDoS dapat membantu mencegah kerusakan yang lebih luas. Penggunaan alat analisis jaringan Wireshark merupakan langkah penting dalam memahami lalu lintas jaringan dan mendeteksi aktivitas mencurigakan(- et al., 2023)..



Gambar 3. Notifikasi hasil Pengujian

Kemampuan model *Decision Tree* untuk mendeteksi serangan DDoS diuji tidak hanya dalam lingkungan simulasi tetapi juga dalam aplikasi nyata, seperti yang terlihat dari sistem peringatan otomatis yang terintegrasi dengan platform komunikasi. Sebagai

contoh nyata dari aplikasi ini, sistem berhasil mendeteksi serangan DDoS. Notifikasi yang dihasilkan, seperti yang ditunjukkan dalam pesan peringatan yang dikirim melalui platform komunikasi, memberikan detail waktu spesifik serangan tersebut terjadi

Penerapan model *Decision Tree* untuk deteksi DDoS menunjukkan pendekatan yang komprehensif dalam menghadapi serangan siber. *Decision Tree* menyediakan mekanisme otomatis untuk deteksi dan respons terhadap serangan tersebut. Keberhasilan dalam mendeteksi serangan dan segera mengirimkan notifikasi memperkuat kapabilitas keamanan siber yang dapat melindungi infrastruktur jaringan dari serangan yang semakin canggih. Efektivitas sistem ini dalam lingkungan operasional nyata menegaskan pentingnya integrasi antara alat analisis teknis dan solusi keamanan berbasis AI, yang secara bersamaan meningkatkan kemampuan deteksi dan respons keamanan jaringan.

SIMPULAN

Model *Decision Tree* yang dikembangkan dalam penelitian ini terbukti efektif dalam analisis real-time dari aliran data TCP, memungkinkan deteksi cepat serangan DDoS. Kegunaan model ini lebih lanjut diperkuat melalui penyimpanan hasil analisis dalam database yang tidak hanya berfungsi sebagai arsip kejadian keamanan tetapi juga sebagai sumber data untuk pelatihan ulang model, memastikan bahwa sistem tetap up-to-date terhadap ancaman terbaru. Selain itu, integrasi sistem dengan platform komunikasi seperti Telegram memungkinkan distribusi peringatan real-time yang memfasilitasi respon cepat terhadap insiden keamanan, meminimalisir potensi kerusakan dari serangan yang terdeteksi.

Berdasarkan hasil dari penelitian ini, beberapa saran untuk pengembangan dan penelitian lanjutan adalah Model *Decision Tree* harus diuji dan mungkin digabungkan dengan algoritma lain seperti jaringan saraf tiruan atau metode ensemble untuk meningkatkan akurasi dan robustness deteksi serangan. Pendekatan hybrid bisa membantu dalam menangani kelemahan yang mungkin ada dalam model tunggal, Selain mendeteksi serangan, sistem harus dikembangkan untuk mengambil tindakan mitigasi otomatis. Ini bisa termasuk otomatisasi pemblokiran IP atau penyesuaian aturan *firewall* berdasarkan jenis serangan yang terdeteksi.

DAFTAR PUSTAKA

-, B. M., -, S. A., -, A. S., & -, R. K. (2023). Exploring Wireshark For Network Traffic Analysis. *International Journal For Multidisciplinary Research*, *5*(6). https://doi.org/10.36948/ijfmr.2023.v05i06.8876

Black, S., & Kim, Y. (2022). An Overview on Detection and Prevention of Application Layer DDoS Attacks. *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*, 0791–0800. https://doi.org/10.1109/CCWC54503.2022.9720741

Budiman, imam. (2022). National Cyber Defense Of The Indonesian Government In Protecting The Society. *Jurnal Mandala Jurnal Ilmu Hubungan Internasional*, 231–243. https://doi.org/10.33822/mjihi.v5i2.4894

Hernández, V. A. S., Monroy, R., Medina-Pérez, M. A., Loyola-González, O., & Herrera, F. (2022). A Practical Tutorial for Decision Tree Induction. *ACM Computing Surveys*, *54*(1), 1–38. https://doi.org/10.1145/3429739

Kowal, D. R. (2022). Fast, Optimal, and Targeted Predictions Using Parameterized Decision Analysis. *Journal of the American Statistical Association*, *117*(540), 1875–1886. https://doi.org/10.1080/01621459.2021.1891926

Ma, Y., Sung, K.-W., & Ahn, H.-J. (2023). N- and F-Co-Doped Carbon Quantum Dots Coated on a Ni Foam Substrate as Current Collector for Highly Stable Li-Air Batteries. *International Journal of Energy Research*, *2023*, 1–11. https://doi.org/10.1155/2023/5310171

P, V., V, P., & K, U. (2021). IMPACTS OF CYBER CRIME ON INTERNET BANKING. *International Journal of Engineering Technology and Management Sciences*, *5*(2). https://doi.org/10.46647/ijetms.2021.v05i02.005

Rao, G. S., & Subbarao, P. K. (2023). A Novel Approach for Detection of DoS / DDoS Attack in Network Environment using Ensemble Machine Learning Model. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(9), 244–253. https://doi.org/10.17762/ijritcc.v11i9.8340

Tedyyana, A., Ghazali, O., & Purbo, O. W. (2024). Machine learning for network defense: automated DDoS detection with telegram notification integration. *Indonesian Journal of Electrical Engineering and Computer Science*, *34*(2), 1102. https://doi.org/10.11591/ijeecs.v34.i2.pp1102-1109

Tsobdjou, L. D., Pierre, S., & Quintero, A. (2022). An Online Entropy-Based DDoS Flooding Attack Detection System With Dynamic Threshold. *IEEE Transactions on Network and Service Management*, *19*(2), 1679–1689. https://doi.org/10.1109/TNSM.2022.3142254

Yacouby, R., & Axman, D. (2020). Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models. *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems*, 79–91. https://doi.org/10.18653/v1/2020.eval4nlp-1.9

Zhang, C., Soda, P., Bi, J., Fan, G., Almpanidis, G., García, S., & Ding, W. (2022). An empirical study on the joint impact of feature selection and data resampling on imbalance classification. *Applied Intelligence*. https://doi.org/10.1007/s10489-022-03772-1