Deteksi Anomali Jaringan Menggunakan *Hybrid Algorithm*

Raden Muhammad Imam[1], Parman Sukarno[2], Muhammad Arief Nugroho[3]

[1,2,3]Fakultas Informatika, Universitas Telkom, Bandung
[1]radenmuhammadimam@students.telkomuniversity.ac.id, [2]psukarno@telkomuniversity.ac.id,
[3]arif.nugroho@telkomuniversity.ac.id

Abstrak

Anomali jaringan adalah suatu keadaan yang terjadi pada sebuah *network traffic* yang menyebabkan kondisi menjadi tidak normal. Untuk mendeteksi anomali jaringan dibutuhkan suatu sistem komputer yang dikenal dengan istilah *Intrusion Detection System* (IDS). Mendeteksi adanya serangan memiliki beberapa kekurangan dan keuntungan. Kekurangan pada *Anomaly-based IDS* yaitu *high false positive* dan *false negative*, sedangkan keuntungan pada *Anomaly-based IDS* yaitu mendeteksi anomali yang dikenal dan tidak dikenal. Pengujian deteksi anomali jaringan menggunakan *hybrid algorithm* dilakukan dengan beberapa skenario untuk mengetahui nilai akurasi dari deteksi anomali yang dihasilkan. Pada tahap analisis dan pengujian, deteksi anomali yang dihasilkan akurasi terbaik sebesar 99.89522%.

Kata kunci : akurasi deteksi, IDS, *hybrid algorithm*, *eager learning*, *lazy learning*

Abstract

Network anomaly is a condition that occurs in a network traffic which causes an abnormal condition. To detect network anomalies we need a computer system known as Intrusion Detection System (IDS). Detecting an attack has several disadvantages and advantages. Lack of Anomaly-based IDS is high false positive and false negative, while the advantage of Anomaly-based IDS is detecting known and unknown anomalies. Testing network anomaly detection using hybrid algorithm is done with several scenarios to determine the accuracy of the resulting anomaly detection. In the analysis and testing phase, the anomaly detection produced the best accuracy of 99.89522%.

Keywords: detection acuraccy, IDS, *hybrid algorithm*, *eager learning*, *lazy learning*

1. Pendahuluan

Pada bagian ini berisi empat sub-bagian, yaitu: Latar Belakang, Topik dan Batasannya, Tujuan dan Organisasi Tulisan. Di bawah ini akan dijelaskan dari masing-masing sub-bagian tersebut.

Latar Belakang

Anomali jaringan adalah suatu keadaan yang terjadi pada sebuah *network traffic* yang menyebabkan kondisi menjadi tidak normal. Anomali yang terjadi bisa dilihat melalui kenaikan lonjakan pengguna Internet, melalui serangan pada suatu *traffic* dan lonjakan yang tidak disengaja. Kenaikan lonjakan dapat dilihat pada saat adanya bencana yang terjadi di dunia, kompetisi atau pertandingan dan kejadian yang tidak biasa terjadi setiap hari. Secara tidak sadar, kondisi kenaikan lonjakan ini memberikan dampak negatif bagi beberapa pihak. Kenaikan lonjakan yang terjadi tersebut menimbulkan penurunan performansi dari suatu jaringan. Untuk itu, perlu dilakukan deteksi terhadap anomali yang terjadi. Untuk mendeteksi anomali jaringan dibutuhkan suatu sistem komputer yang dikenal dengan istilah *Intrusion Detection System* (IDS)[28].

Intrusion Detection Systems (IDS) adalah sistem pertahanan dan keamanan otomatis untuk *monitor*, mendeteksi, dan menganalisis *hostile activities* dalam jaringan atau host[18]. *Firewall* biasanya adalah garis pertahanan pertama dalam jaringan dan IDS digunakan ketika ada bukti anomali, yang tidak dapat dihentikan atau dikurangi oleh *firewall*. IDS kemudian berfungsi sebagai garis pertahanan kedua. Selain itu, tugasnya sulit dan pada kenyataannya sistem deteksi anomali tidak mendeteksi intrusi sama sekali, hanya mengidentifikasi bukti intrusi, baik ketika sedang berlangsung atau setelah fakta IDS dapat dikategorikan dalam banyak cara[5]. Selanjutnya, mengenai teknik mendeteksi aktivitas yang tidak biasa, IDS dapat dikategorikan menjadi empat jenis: *Anomaly-based IDS*, *Signature-based IDS*, *Specification-based IDS* dan *Hybrid*[16].

Mendeteksi adanya serangan memiliki beberapa kekurangan. Kekurangan pada *Anomaly-based IDS* yaitu *high false positive* dan *false negative*; *Signature-based IDS* yaitu tidak dapat mendeteksi anomali yang tidak dikenal,

difficult dan *time-consuming task* untuk *build* dan *update signatures*; *Specification-based IDS* yaitu elaborasi spesifikasi terperinci dan kendala memakan biaya dan waktu, sedangkan *Hybrid IDS* yaitu terbatas pada operasi yang tepat dari suatu program atau protokol. Mendapatkan pendekatan berbeda untuk *interoperate* dan *coexist* dalam *single system*.

Pada penelitian ini, sebuah algoritma *hybrid* dari dua jenis *supervised learning* yaitu *Eager Learning* dan *Lazy Learning*. Untuk *eager learning*, digunakan *rule induction* (dengan algoritma RIPPER), sedangkan untuk *lazy learning* digunakan *instance-based learning* (dengan algoritma *K-Nearest Neighbour*). *Rule induction* (dengan algoritma RIPPER) memiliki kinerja yang unggul berdasarkan kondisi dimana terdapat jumlah atribut yang tinggi dan kemampuan pemrosesan yang tinggi. Namun, kekurangan dari algoritma RIPPER ini adalah kinerjanya akan menurun ketika diberikan *training set* dengan jumlah yang besar[29].

Sedangkan pendeteksian intrusi dengan algoritma *K-Nearest Neighbour* mempunyai kelebihan dapat mendeteksi intrusi yang sebelumnya tidak dikenali. Dengan menggabungkan keduanya, menghasilkan *machine learning* yang mempunyai kinerja yang baik pada tahap *learning* maupun tahap klasifikasi serta mempunyai tingkat pendeteksian intrusi yang tinggi baik untuk intrusi yang sudah pernah dipelajari maupun yang sebelumnya tidak dikenali.

Penelitian dengan menggunakan kedua algoritma ini sudah pernah dilakukan sebelumnya[29], namun penggunaan *dataset* yang berbeda. Dengan menggabungkan keduanya, akan dihasilkan *machine learning* yang mempunyai kinerja yang baik dalam fase *learning* maupun fase klasifikasi serta mempunyai tingkat pendeteksian intrusi yang tinggi, baik untuk intrusi yang sudah pernah dipelajari maupun yang sebelumnya tidak dikenali.

Topik dan Batasannya

Pada penelitian ini, menganalisa data trafik jaringan yang digunakan bersifat *offline*, yaitu data NSL-KDD berdasarkan parameter Akurasi. Dalam menganalisa data trafik tersebut, menggunakan *Intrusion Detection System* (IDS). Dalam membangun IDS menggunakan algoritma *hybrid*. Algoritma *hybrid* terdiri dari *eager learning* dan *lazy learning*. Algoritma *eager learning* yang digunakan adalah algoritma RIPPER, sedangkan algoritma *lazy learning* yang digunakan adalah algoritma *K-Nearest Neighbour*.

Tujuan

Penelitian ini dilakukan dengan tujuan untuk mendeteksi anomali dengan menggunakan *hybrid algorithm* dalam pembuatan *Intrusion Detection System* (IDS). Penelitian ini juga bertujuan untuk menganalisis IDS yang dihasilkan berdasarkan parameter Akurasi.

Organisasi Tulisan

Selanjutnya, penelitian ini akan menjelaskan berbagai penelitian yang sudah pernah dilakukan sebelumnya serta hal-hal yang terkait dengan penelitian ini pada bagian II. Metode penelitian yang digunakan dan sistem yang dibangun pada proses klasifikasi akan dijelaskan pada bagian III. Pada bagian IV akan dijelaskan hasil yang diperoleh serta evaluasi dari penelitian yang telah dilakukan. Terakhir yaitu penarikan kesimpulan dan pemberian saran untuk penelitian kedepannya yang akan dijelaskan pada bagian V.

2. Studi Terkait

2.1 Anomali Jaringan

Anomali jaringan adalah suatu keadaan yang terjadi pada sebuah *network traffic* yang menyebabkan kondisi menjadi tidak normal. Anomali yang terjadi bisa dilihat melalui kenaikan lonjakan pengguna Internet, melalui serangan pada suatu *traffic* dan lonjakan yang tidak disengaja. Kenaikan lonjakan dapat dilihat pada saat adanya bencana yang terjadi di dunia, kompetisi atau pertandingan dan kejadian yang tidak biasa terjadi setiap hari. Secara tidak sadar, kondisi kenaikan lonjakan ini memberikan dampak negatif bagi beberapa pihak. Kenaikan lonjakan yang terjadi tersebut menimbulkan penurunan performansi dari suatu jaringan[28]. Anomali jaringan dapat dikategorikan memberikan dua sifat yang relevan: anomali berdasarkan sifatnya dan anomali berdasarkan aspek penyebabnya[16].

2.1.1 Anomali Berdasarkan Sifatnya

Sifat anomali adalah aspek penting dari teknik deteksi anomali. Aspek ini dapat mengarahkan bagaimana sistem akan menangani dan memahami anomali yang ditambang dan dideteksi. Berdasarkan sifatnya, ada tiga kategori anomali: *point anomalies*, *collective anomalies* dan *contextual anomalies*[8],[4],[1].