Peningkatan Kinerja pada Intrusion Detection System (IDS) dengan 3 Lapisan Pembelajaran Hibrid = Improved Performance in Intrusion Detection System (IDS) with 3 Layer Hybrid Learning

Fajar Henri Erasmus Ndolu, author

Abstrak

Dengan perkembangan teknologi informasi yang pesat saat ini, serangan siber terhadap jaringan semakin meningkat dan menyebabkan kerugian finansial yang signifikan. Oleh karena itu, sistem deteksi intrusi (IDS) berbasis anomali menggunakan pembelajaran mesin menjadi salah satu pendekatan untuk mendeteksi serangan siber. Tetapi, penggunaan algoritma tunggal dalam IDS memiliki kekurangan dalam mendeteksi jenis serangan yang memiliki kelas minoritas dalam dataset. Selain itu, penggunaan dataset yang tidak seimbang dan tidak mencerminkan kondisi saat ini juga mempengaruhi kinerja IDS. Untuk meningkatkan kinerja IDS, diusulkan metode hibrid dengan menggunakan Long Short Term Memory (LSTM) dan Random Forest (RF), dengan dataset terbaru CIC-CSE-IDS2018. Dalam pembentukan model hibrid, model lapisan satu menggunakan LSTM untuk klasifikasi biner, mengklasifikasikan aliran data sebagai data normal atau data serangan. Data normal diklasifikasikan kembali dengan model lapisan dua dan data serangan diklasifikasikan kembali dengan model lapisan tiga. Jika hasil model lapisan dua diklasifikasikan sebagai data normal, maka merupakan hasil akhir, dan jika diklasifikasikan sebagai data serangan maka diklasifikasikan kembali dengan model lapisan tiga secara multikelas menggunakan RF. Hasil klasifikasi multikelas lapisan tiga merupakan hasil akhir dari model hibrid ini. Berdasarkan pengujian dan analisis, model hibrid dengan evaluasi terbaik di peroleh menggunakan dataset dengan rasio 3 : 1. Model hibrid ini mencapai hasil klasifikasi multi kelas dengan accuracy 99,7618%, precision 99,1901%, recall 96,8809% dan f1-score 97,9508%.

.....With today's rapid development of information technology, cyber attacks against networks are increasing and causing significant financial losses. Therefore, an anomaly-based intrusion detection system (IDS) using machine learning is one approach to detecting cyber attacks. However, the use of a single algorithm in IDS has drawbacks in detecting types of attacks that have a minority class in the dataset. In addition, the use of unbalanced datasets that do not reflect current conditions also affects IDS performance. To improve IDS performance, a hybrid method is proposed using Long Short Term Memory (LSTM) and Random Forest (RF), with the latest CIC-CSE-IDS2018 dataset. In the hybrid model, the layer one model uses LSTM for binary classification, classifying the data stream as normal data or attack data. Normal data is reclassified by layer two model and attack data is reclassified by layer three model. If the result of the second layer model is classified as normal data, then it is the final result, and if it is classified as attack data then it is reclassified with the third layer model in a multiclass manner using RF. The results of the three layer multiclass classification are the final results of this hybrid model. Based on testing and analysis, the hybrid model with the best evaluation was obtained using a dataset with a ratio of 3:1. This hybrid model achieved multiclass classification results with 99.7618% accuracy, 99.1901% precision, 96.8809% recall and f1-score 97.9508%.