Analisis Perbandingan Kinerja Metode Seleksi Fitur Information Gain Ratio dan Chi-Square dalam Klasifikasi Serangan Siber pada Jaringan Wi-Fi = Comparative Performance Analysis Of Information Gain Ratio and Chi Square Feature Selection Methods in Cyber Attack Classification On Wi-Fi Networks

Nurul Qomariah Abdillah, author

Abstrak

Perkembangan teknologi informasi dan komunikasi saat ini menciptakan ketergantungan manusia terhadap teknologi dan internet, salah satunya melalui penggunaan jaringan Wi-Fi. Konektivitas Wi-Fi berkaitan erat dengan Internet of Things (IoT) karena dapat memfasilitasi perangkat IoT untuk saling terhubung dan terkoneksi ke jaringan internet. Namun, peningkatan penggunaan Wi-Fi publik maupun privat rentan terhadap serangan siber. Badan Sandi dan Siber Negara memperkirakan tahun 2024 akan muncul ancaman seperti IoT attacks, distributed denial of services (DDOS), phishing, dan lainnya. Oleh karena itu, perlu adanya upaya antisipatif untuk mengatasi serangan siber. Salah satu upayanya adalah menerapkan intrusion detection system (IDS) untuk memantau lalu lintas jaringan dan memberikan peringatan jika terdapat serangan. Peningkatan kemampuan deteksi IDS dapat dilakukan dengan menerapkan metode machine learning yang mampu mempelajari pola serangan secara efektif dan akurat. Pada penelitian skripsi ini diterapkan metode klasifikasi Support Vector Machine (SVM) Multiclass dengan pendekatan one-vs-one dan one-vs-rest pada dataset Aegean Wi-Fi Intrusion Detection System (AWID2) yang terdiri dari empat kelas dan memiliki dimensi data yang tinggi, yaitu 154 dimensi (fitur). Dalam mengatasi masalah dimensi tinggi tersebut dilakukan seleksi fitur yang bertujuan untuk menghilangkan fitur yang tidak relevan, sehingga fitur hanya terkonsentrasi pada fitur- fitur yang relevan dan informatif dalam menggambarkan serangan. Penelitian skripsi ini menggunakan metode Chi-square dan Information Gain Ratio. Hasil penelitian skripsi ini menunjukkan metode seleksi fitur Chi-square dengan klasifikasi SVM One Vs Rest pada kernel polynomial dengan memilih 54 fitur tertinggi merupakan model terbaik dalam mengklasifikasikan serangan siber pada Wi-Fi dengan nilai accuracy = 98,03%, Precision = 87,24%, Recall = 99,30%, dan F1 score = 91,90%.

.....Today's advances in information and communication technology create human dependence on technology and the Internet, one of which is through the use of Wi-Fi networks. Wi-fi connectivity is closely related to the Internet of Things (IoT) because it can facilitate IoT devices to interconnect and be connected to the internet network. However, increased use of public and private Wi-FI is vulnerable to cyber attacks. The National Password and Cyber Agency predicts that threats such as IoT attacks, Distributed Denial of Services, phishing, and more will emerge in 2024. Therefore, there is a need for pre-emptive efforts to deal with cyberattacks. One attempt is to implement the Intrusion Detection System (IDS) to monitor network traffic and give warning if there is an attack. Improved IDS detection capabilities can be achieved by applying machine learning methods that can learn patterns of attack effectively and accurately. In this study, the multi-class Support Vector Machine (SVM) classification method was applied to the Aegean Wi-Fi Intrusion Detection System (AWID2) dataset, which consists of four classes and has a high data dimension, namely 154 dimensions. In addressing the high dimension problem, a feature selection was carried out

aimed at eliminating irrelevant features, so that the features were concentrated only on the features that are relevant and informative in describing the attack. This study of the script uses the Chi-square method and Information Gain Ratio. The results of this study show that the method of selection of the feature Chi-square with SVM One vs Rest classification on the polynomial kernel by choosing the 54 highest features is the best model in classifying cyber attacks on Wi-Fi with accuracy values = 98.03%, Precision = 87.24%, Recall = 99.30%, and F1 score = 91.90%.