ABSTRAK

Meningkatnya frekuensi serangan *ransomware*, serta munculnya variasi-variasi baru, merupakan bahaya besar bagi keamanan siber. Email *phishing*, perangkat lunak berbahaya, unduhan ilegal, dan eksploitasi kerentanan sistem adalah metode umum untuk menyebarkan *ransomware*. Penelitian pendeteksian *ransomware* sering kali menggunakan analisis statis dan dinamis, tetapi kendala yang paling sulit adalah mengidentifikasi serangan *zero-day* dan waktu yang dibutuhkan untuk melakukannya. Analisis dinamis, meskipun membutuhkan eksekusi serangan, seringkali menghasilkan temuan yang lebih akurat daripada analisis statis. Penelitian ini mencoba untuk meningkatkan deteksi dan pencegahan *ransomware* secara *response-time* sebelum file dienkripsi pada sistem *host*. Metode deteksi berbasis *host* digunakan, termasuk teknik *windowing* dan ekstraksi fitur menggunakan n-gram dan TF-IDF. Model pendeteksian dibuat dengan menggunakan tiga algoritma pembelajaran mesin: Naive Bayes, Random Forest, dan Support Vector Machine. Kriteria yang digunakan dalam evaluasi adalah *accuracy, precision, recall, F1-score*, *false positive rate* (FPR), dan *false negative rate* (FNR). Hasilnya menunjukkan bahwa Naive Bayes dengan 8 dan 10 n-gram mendapatkan *accuracy, precision, recall, F1-score* sebesar 100%, dengan FPR dan FNR sebesar 0%, yang lebih baik daripada model lainnya. Selain itu, model Naive Bayes dengan n-gram antara 2 dan 10 memberikan hasil matriks evaluasi terbaik dan paling konsisten setelah lebih dari 20.000 pemanggilan API. Model-model lainnya memberikan hasil terbaik setelah lebih dari 30.000 panggilan API. Sistem yang disarankan dapat mengidentifikasi *ransomware* pada kisaran 25.000 hingga 35.000 panggilan API. Penelitian ini juga menemukan variasi fase antara aktivitas normal dan *ransomware*, terutama pada tahap awal dan akhir. Meskipun *ransomware* dan aktivitas biasa memiliki periode inisialisasi yang sama, fase *ransomware* berbeda secara signifikan, mendukung deteksi pada tahap awal. Dengan demikian, deteksi pada kisaran 25.000 panggilan API sangat membantu dalam mencegah kesalahan identifikasi *ransomware* sebagai aktivitas biasa. Akan tetapi, meskipun menggunakan n-gram dalam jumlah besar meningkatkan *accuracy*, hal itu juga meningkatkan ukuran data dan waktu eksekusi. Akibatnya, memilih n-gram yang tidak terlalu tinggi atau terlalu rendah sangat penting untuk mempertahankan kinerja yang seimbang. *Framework* ini juga memiliki banyak pengaturan, termasuk ukuran jendela, ambang deteksi, dan pengulangan sepuluh jendela. Penggunaan n-gram pada kerangka kerja *response-time* ini telah terbukti berguna dalam mengidentifikasi *ransomware* sebelum mengenkripsi semua data.

Kata kunci: Deteksi *ransomware*, *response-time*, *windowing*, berbasis *host*, serangan *zero-day*, panggilan API