Analisis big data penyebaran Malware Avalanche Berbasis Clustering dengan Algoritma K-means pada Infrastruktur Internet Indonesia = Big data analysis of Clustering-Based Avalanche Malware Spread with K-means Algorithm on Indonesia's Internet Infrastructure

Nidaul Muiz Aufa, author

Abstrak

Tesis ini membahas penyebaran malware Avalanche pada infrastruktur internet Indonesia. Penelitian dilakukan dengan metode analisis big data dengan menggunakan Algoritma K-mean (k=3). Dataset pada penelitian ini menggunakan dataset yang diperoleh dari CERT-bund. Hasil penelitian ini menggambarkan bahwa infrastruktur internet Indonesia masih terinfeksi malware Avalanche dengan aktivitas sebanyak 44.254.374 sepanjang tahun 2018 dan 2019. Aktivitas ini melibatkan 969 AS Number, 3.173.254 IP Address, dan 26 jenis malware. Hasil Clustering menggunakan Splunk terhadap AS Number dan IP Address menghasilkan masing-masing 3 cluster. Cluster AS Number yang paling produktif adalah cluster1 yang memiliki populasi 3 AS Number. Sedangkan Cluster IP Address yang paling produktif adalah cluster1 dengan populasi 32.991 IP Address.

.....This thesis discusses the spread of Avalanche malware on Indonesian internet infrastructure. The research was conducted by using the big data analysis method using the K-mean algorithm (k = 3). The dataset in this study was obtained from the CERT-bund. The results of this study illustrate that Indonesia's cyber infrastructure is still infected with Avalanche malware with a total of 44,254,374 activities throughout 2018 and 2019. This activity involved 969 AS Numbers, 3,173,254 IP Addresses, and 26 types of malware. The results of clustering using Splunk on the AS Number and IP Address resulted in 3 clusters each. The most productive AS Number cluster is cluster1 which has a population of 3 AS Number. Meanwhile, the most productive cluster IP address is cluster1 with a population of 32,991 IP addresses.