

Abstract

Information in website log data is very important for monitoring web servers. Web servers in the campus environment have an important role in learning, but at this time the lack of attention to web server network security results in the vulnerability of web servers to interference from unauthorized parties. The Security Information and Event Management (SIEM) system is used to monitor the activities of visitors to the campus website but cannot detect anomalies. This final project adds the SIEM function of detecting anomalous patterns. The SIEM proposed in this final project can detect anomaly patterns using information from web server logs. The traffic anomaly detection method used is the Clustering method. K-Means Clustering algorithm is a method of grouping data based on similarity of attributes to form clusters. DBSCAN (Density Based Spatial Clustering) is a density-based clustering algorithm with Noise. The web server log data used as Clustering attributes are Status Code, URL, and Response Size, URL, and Status Code. Attribute similarity refers to identifying similar patterns in the magnitude of Status Code, URL, and Response Size. After the data is clustered, data that has a significant distance from the cluster center and outlier data are considered as anomalies. Tests were conducted with log data from the bis-sby.telkomuniversity.ac.id website with 7 days of data collection totaling 21,892 data. The results of anomaly pattern detection on web server logs include Status Code, URL, and Response Size which are above the average of web server log data for 7 days with an average value of 5,846 bytes in campus website access patterns.

Keywords: *log web server, k-means clustering, DBSCAN, anomaly.*