ABSTRACT

Technological advances and the increasingly rapid development of the internet have resulted in more information and data that needs to be protected because there are many ways hackers can get information or data. The vulnerability of a network is also caused by the more open knowledge about hacking. Distributed Denial of Service (DDoS) is an attack with more than one attacker flooding the packets to the server, so that the server is busy serving very many packet requests and makes server performance decrease. Much research has been done to detect DDoS attacks. However, the research carried out still uses old datasets that have not kept up with the development of DDoS attack trends. This final project classifies the CICIDS2018 DDoS attack dataset using the Naïve Bayes algorithm and Random Forest. Characteristics of the dataset that have been extracted also need to be selected for features in order to shorten the time for data training so that the process increases the efficiency of the classification algorithm. In this final project, feature selection is carried out using the Information Gain method to look for features that have a big influence in determining whether a packet sent is a DDoS attack or not. The use of the Information Gain method to perform feature selection on the CICIDS2018 dataset produces six optimal features, including src_ip, dst_ip, flow_duration, flow_iat_max, fwd_iat_max, and bwd_iat_tot. The test results show the level of DDoS attack detection accuracy for the Naïve Bayes algorithm is 69.6% and for the Random Forest algorithm is 97.2%.

Keywords: Distributed Denial of Service (DDoS), Information Gain, Naïve Bayes, Random Forest.