

ABSTRAK

Karena meningkatnya volume dan keragaman lalu lintas jaringan, sistem deteksi intrusi tradisional (IDS) menjadi tidak memadai akibat kurangnya data berlabel, ketidakseimbangan kelas yang parah, dan kekakuan ambang batas kepercayaan dalam pembelajaran semi-supervised. Untuk mengatasi tantangan ini, penelitian ini mengusulkan kerangka kerja co-training bertahap yang memanfaatkan algoritma Naive Bayes dan Random Forest dengan propagasi label berbasis graf. Mengakui bahwa ambang batas kepercayaan statis dapat memperkenalkan label yang berisik atau membatasi pembelajaran secara berlebihan, ambang batas adaptif yang dihasilkan dari varian Flexible Confidence (FlexCon) diterapkan untuk mengatur inklusi label palsu. Selain itu, tiga representasi fitur—asli, IPCA, dan proyeksi acak—dibandingkan untuk memastikan keseimbangan antara akurasi deteksi dan beban komputasi. Dengan memanfaatkan data berlabel dan tidak berlabel, desain ini menyeimbangkan ketidakseimbangan kelas sambil secara dinamis menyesuaikan kriteria penandaan untuk meningkatkan kinerja deteksi anomali multi-kelas pada sistem deteksi intrusi berskala besar.

Kata kunci: Semi-Supervised, CICIoT2023, FlexCon, Co-Training, Pengurangan Dimensi.