Enhanced intrusion detection capabilities via weighted chi-square, discretization and SVM

ABSTRACT

Anomaly Intrusion Detection Systems (ADSs) identify patterns of network data behaviour to determine whether they are normal or represent an attack using the learning detection model. Much research has been conducted on enhancing ADSs particularly in the area of data mining that focuses on intrusive behaviour detection. Unfortunately, the current detection models such as the support vector machine (SVM) is affected by high dimensional data which limits its ability to accurately classify data. Moreover, the data points which appear similar between intrusive and regular behaviours could be problematic as some innovated attack behaviours may not be detected. To overcome this SVM drawback, we propose a combination of weighted chi-square (WCS) as a feature selection (FS) and a Discretization process (D). The WCS method is used firstly to reduce the dimensionality of data following which the assembled records are transformed into interval values via the D process before the SVM is used to identify groups of samples that behave similarly and dissimilarly such as malicious and non-malicious activities. Experiments were performed with well-known NSL-KDD data sets and the results show that the proposed method namely WCS-D-SVM (weighted chi-square, discretization and support vector machine) significantly improved and enhanced accuracy and detection rates while decreasing the false positives which the single SVM classifier produces.

Keyword: Intrusion detection; Data mining; Feature selection; Weighted chi-square, Discretization; Support vector machine