Data security using unsupervised learning and explanations

Vulnerability assessment is an effective security mechanism to identify vulnerabilities in systems or networks before they are exploited. However manual analysis of network testing and vulnerability assessment results is time consuming and demands expertise. This paper presents an improvement of Analia, which is a security system to process results obtained after a vulnerability assessment using artificial intelligence techniques. The system applies unsupervised clustering techniques to discover hidden patterns and extract abnormal device behaviours by clustering devices in groups that share similar vulnerabilities. The proposed improvement consists in extracting a symbolic explanation for each cluster to help security analysts to understand the clustering solution using network security lexicon.