Abstract

This paper describes a mechanism for detection of changes in a data set to be protected or monitored. The mechanism is based on the creation of detectors derived from self character strings from the data set. The data structure used for storing detectors is a tree array. Each tree in the array stores detectors for a sequence of r adjacent characters within a character string extracted from the data set. Each node in a tree represents a character in a detector. In the monitoring phase, changes in the data set are detected going across nodes of the tree array. A character string being analyzed is compared with detectors in the tree array. A change is detected when no matching route from the root to its leaves is found. Effectiveness in detector generation and change detection of the proposed mechanism is tested with experiments consisting of generating all possible character chains for a binary alphabet and executing a detection or monitoring phase, that is, classifying a set of character strings as proper or nonproper. The best nonproper string detection results were obtained when r is equal to 1.

Keywords