

ABSTRACT

A method and a system for clustering executable files are provided. The method comprises: obtaining a plurality of executable files; for each executable file: (i) detecting repeat sequences of commands of a predetermined length in a given executable file; (ii) determining at least one frequently occurring sequence of the repeat sequences in the given executable file; and based on the at least one frequently occurring sequence of commands, attributing the given executable file to a respective family; iteratively executing the detecting, the determining, and the attributing until one of: all of the plurality of executable files are attributed to at least one respective family, and until un-attributed files of the plurality of executable files do not contain any repeat sequences of commands; and responsive to presence of un-attributed files, attributing each of the un-attributed files of the plurality of executable files to a separate family.

Fig. 1A