

METHOD AND APPARATUS FOR IDENTIFYING APPLETS OF RISKY CONTENT BASED ON DIFFERENTIAL PRIVACY PRESERVING

ABSTRACT

Implementations of the present specification provide a method and an apparatus for identifying an applet of risky content based on differential privacy preserving. A specific implementation of the method includes: in response to monitoring a target operation performed by a user by using an applet carried by a service platform, obtaining a result of determining whether an operation source is a risky operation source, the operation source including a combination of the user and a device used by the user, and the determining result being determined by using a risk identification policy group; in response to determining that the operation source is a risky operation source, disturbing a program identifier of the applet by using a randomized response technique that satisfies local differential privacy preserving, to obtain a disturbed program identifier; and sending an operation source identifier of the operation source and the disturbed program identifier to a serving end of the service platform, so that the serving end identifies whether the applet is an applet of risky content based on multiple received disturbed program identifiers by using the randomized response technique.

FIG. 2