Bo YU, Ying FANG, Qiang YANG, Yong TANG, Liu LIU, 2018. A survey of malware behavior description and analysis. *Frontiers of Information Technology & Electronic Engineering*, 19(6):583-603. https://doi.org/10.1631/FITEE.1601745

A survey of malware behavior description and analysis

Key words: Malware behavior; Static analysis; Dynamic Analysis; Behavior data expression; Behavior analysis; Machine learning; Semantics-based analysis; Behavior visualization; Malware evolution

Corresponding author: Bo YU
E-mail: yubo0615@nudt.edu.cn
ORCID: http://orcid.org/0000-0001-6576-5555

Motivations

1. The quantity and complexity of malware samples have increased considerably over the past few years.

2. Behavior-based malware analysis is an important technique for automatically analyzing and detecting malware, and it has received considerable attention from both academic and industrial communities

3. Existing dynamic approaches help analysts understand the behavior intentions and evolution trends.

4. We conduct a survey on malware behavior description and analysis considering three aspects: malware behavior description, behavior analysis methods, and visualization techniques.

Main ideas

1. Existing behavior data types and emerging techniques for malware behavior description are explored, especially the goals, principles, characteristics, and classifications of behavior analysis techniques proposed in the existing approaches.

2. The inadequacies and challenges in malware behavior analysis are summarized from different perspectives.

3. Several possible directions are discussed for future research.

Methods

1.Malware behavior expression discussion, including behavior data types, malware behavior description.

2. Behavior analysis method, including machine learning based behavior analysis, semantics-based behavior identification ,and hybrid approach.

3. Visualization techniques of behavior analysis, existing visualization techniques, such as the similarity map, sequence graph, and tree-map, are used to explore behavior relationships between malware samples, aiding malware analysis in a visual way.

Analysis and discussions

We will discuss the categorization of related works from several perspectives:

(1) the goal of the work;

(2) what kind of behavior data are obtained;

(3) and which analysis techniques are used for a specific behavior-analysis goal.

Hence, the classification of existing works includes:

(1) classification of analysis goals;

(2) classification of behavior data levels;

(3) classification of analysis techniques.

Challenges and suggestions

We can summarize several considerable challenges and suggestions. Although there are many works on malware behavior analysis, more efforts are still needed for better effectiveness and accuracy. The challenges under consideration are as follows:

(1) coverage of behavior data;
(2) unknown behavior detection;
(3) malware adversarial behavior;
(4) malware evolution analysis.

Conclusions

1. We have performed a comprehensive review of the latest malware behavior analysis techniques and discussed the existing research classified from five different perspectives, clearly showing the advantages and disadvantages of existing analysis methods.

2. We have discussed some inadequacies and challenges that are currently not solved as well as several possible solutions to address the current shortcomings.