

Malware Analysis

Provide Detailed Analysis of Malware Code Using A Combination of Static and Dynamic Techniques and Tools

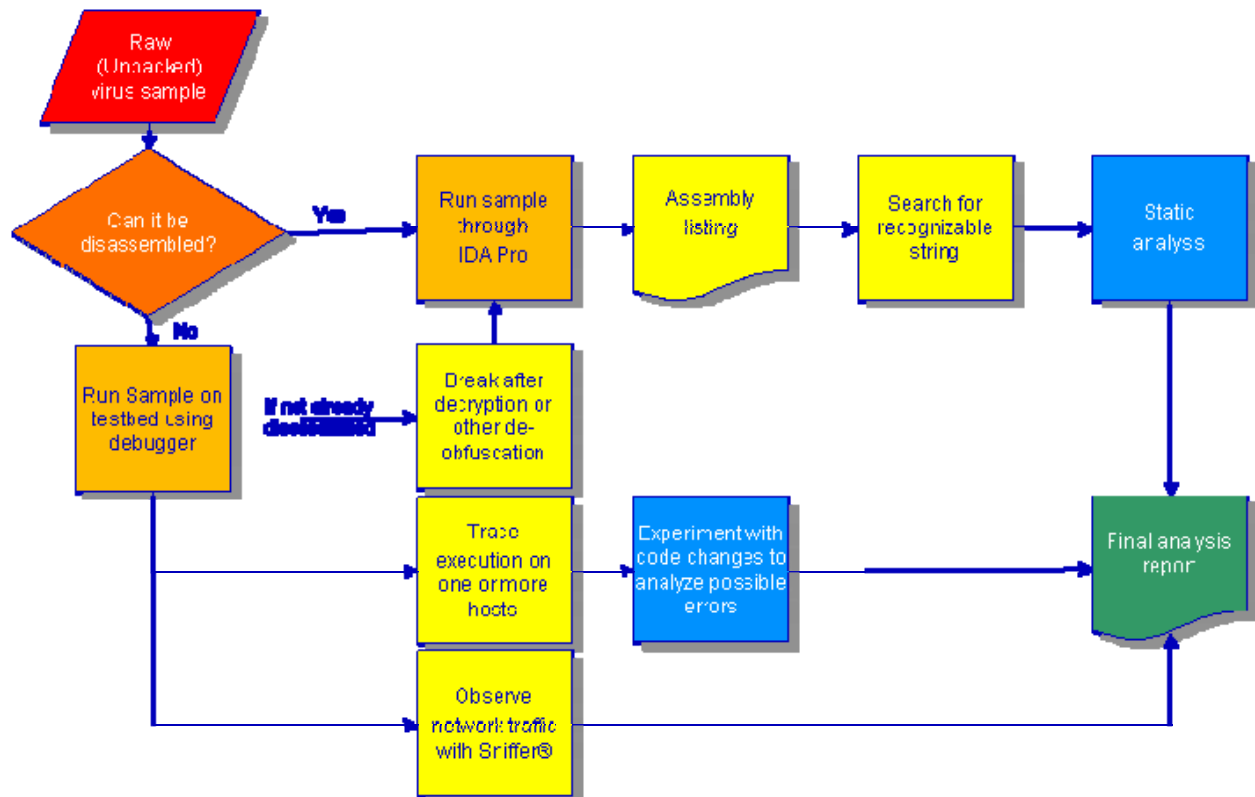
SPARTA analyzes malware samples to provide in-depth understanding of individual samples and to identify emerging technical trends from large collections of malware samples. Nearly all of the malware samples are Windows-compatible binary executables (without source code). In-depth analysis of single malware samples provides information on:

- Capabilities
- Construction techniques
- Anti-reverse engineering techniques
- Similarity to other samples

SPARTA employs a combination of static and dynamic analysis techniques as part of a unified malware analysis process. The majority of malware samples require unpacking as the initial

step in the process in order to get an uncompressed or decrypted form of the malware sample. The next steps involve multiple iterations of static disassembly, interactive debugging, and isolated execution with host and network monitoring. The number of cycles and total time for analysis varies with the complexity of the sample. The figure below shows the cycle of steps involved in the process.

Static analysis focuses specifically on determining a sample's behavior by exploring the characteristics of the sample's code. This form of analysis aids in determining characteristics of sample that are not necessarily observable from performing dynamic analysis alone. For example, a virus may use code obfuscation techniques, such as encryption or entry point obscuration, to make debugging



Malware Analysis Process

difficult. The malware sample may also not exercise all possible execution paths during dynamic analysis.

Dynamic analysis monitors malware while they are executing. We use monitoring tools to observe how the malware interacts with host system resources (e.g. file system and system registry) and generates network traffic in the test environment. The test environment simulates internet resources that the malware sample may require for execution. We then evaluate the data points to determine the malware's functionality and possible objectives.

Static and dynamic analysis offer different, but complementary information about a malware sample. For example, a sample's use of encryption may prevent us from initially observing the code statically. Using dynamic techniques, we can extract the decrypted code to observe it statically. As another example, dynamic analysis may not reveal that a sample attempts to implement a genetic algorithm, while we may readily observe this feature statically.

SPARTA also analyzes large collections of malware samples to identify changes in technical trends, either in malware capabilities or

construction techniques. SPARTA's collection includes tens of thousands of samples and is constantly growing as we collect more samples from the wild. The figure below shows the significant steps in the process. As with in-depth analysis, the initial step is to unpack the malware samples. Our analysis process is based on identifying patterns of text strings in the sample that are indicative of larger malware features. Examples of such features are reading a file, spawning threads, or using IRC. We record the malware samples, strings, and feature information in a relational database from which we can generate periodic trends reports.

We also use the large-scale feature analysis process to identify samples that may deserve in-depth analysis. We need to separate the "interesting" samples from the large number of common variants that are found every day. To do this, we apply data mining tools to the larger collection. The data mining tools partitioned the collection into clusters based upon selected clustering algorithms. We then analyze these clusters to look for samples of interests. These could be malware samples that appear as outliers within their own small cluster or samples that form sub-clusters within larger clusters.

Malware Profiling Process

