Role of Computational Intelligence
in
Malware Detection and Malicious Code Analysis

Srinivas Mukkamala Institute for Complex Additive Systems Analysis

Computational Analysis and Network Enterprise Solutions

New Mexico Tech

Cyber Security Works

smukkamala@caanes.com | srinivas@cs.nmt.edu

Proliferation of malicious code (viruses, worms, Trojans, rootkits, spyware, crime ware, phishing attacks, and other malware designed to infiltrate or damage a system without user's consent) *in recent years has presented a serious threat to Internet, individual users, and enterprises alike.* Current static scanning techniques for malicious code detection have serious limitations; on the other hand, sandbox testing fails to provide a complete satisfactory solution either due to time constraints (e.g., time bombs cannot be detected before its preset time expires).

What is making the situation worse is the ease of producing polymorphic (or variants of) and metamorphic computer viruses that are even more complex and difficult than their original versions to detect. To make the situation even worse malware once confined to wired networks, has now found a new breeding ground in mobile devices, automatic identification and collection (AIDC) technologies, and radio frequency identification devices (RFID) that use wireless networks to communicate and connect to the Internet.

In order to stay ahead and be proactive in an asymmetric race against malicious code writers, developers of antimalware technologies have to rely on automatic malware analysis tools. The hypothesis is that all versions of the same malicious code or similar malware share a common core signature that is a combination of several features of the code. After a particular malicious code has been first identified, it can be analyzed to extract the signature, which provides a basis for detecting variants and mutants of the same malware in the future.

In this Talk, we introduce a method of functionally classifying malicious code and intrusions by using well-known computational intelligent techniques. We present the performance of different methods and describe results of kernel methods in the context of classification accuracy on malware datasets, well know intrusion detection, phishing, denial of service attacks datasets. Our results thus demonstrate the potential of using computational intelligent techniques for detecting and classifying malicious programs from benign programs.

We also discuss other application areas where computational intelligent techniques can be applied for lung cancer detection based on same core principles used for detecting malicious programs.