Chapter X
Limitations of Current Anti-Virus Scanning Technologies

Srinivas Mukkamala
New Mexico Tech, USA

Antonins Sulaiman
New Mexico Tech, USA

Patrick Chavez
New Mexico Tech, USA

Andrew H. Sung
New Mexico Tech, USA

ABSTRACT

Malware has become more lethal by using multiple attack vectors to exploit both known and unknown vulnerabilities and can attack prescanned targets with lightning speed. In the future, it is important that the scanners are capable of detecting polymorphic (obfuscated or variant) and metamorphic (mutated or evolved) versions of malware, however current scanning techniques for malware detection have serious limitations. Simple software obfuscation a general technique that is used to protect the software from reverse engineering techniques can circumvent the current detection mechanisms (anti-virus tools). In this chapter, we describe common attacks on anti-virus tools and a few obfuscation techniques applied to recent viruses that were used to thwart commercial grade anti-virus tools. Similarities among different malware and their variants are also presented in this chapter. The signature used in this method is the percentage of application programming interface (APIs) appearing in the malware type. The hypothesis is that mutants and variants will not stray far from the original. Table 5 shows serious limitations of commercial grade anti-virus scanners in detecting simple obfuscation attacks. Table 6 shows the percentages of similarity of a particular malware when compared to others. One important thing to note is that even the polymorphic ZMist uses the same set of APIs on all three variants.
INTRODUCTION

The circle of attack and defense in the world of malicious software is one that never ends. Anti-virus companies are competing to devise their best scanning technology, while the malware writers are devising every possible way to defeat the scanners. So the war against malware continues.

Internet worms, Trojans, and backdoors are now a significant growing threat, alongside EXE infectors and macro-viruses. Increasingly, the term malware is used to encompass all threats. A malware or malicious code is a piece of code that can affect the secrecy, integrity, data and control flow, and functionality of a system. Therefore, detection is a major concern within the research community as well as within the user community. As malicious code can affect the data and control flow of a program, static analysis may naturally be helpful as part of the detection process (Christodorescu & Jha, 2003).

Malicious software is classified broadly based on the payload and propagation mechanism. In this work, we are classifying malware based on their behavioral pattern. By doing this, we can use our techniques, which are based on similarities to the known malware signature. The main goal here is to find a similar pattern for each class that we can use to identify future malware based on that class. So our next section will briefly introduce the malware classifications that we use for our purpose.

Malware categorization based on the behavioral patterns is described further in this chapter. Considering that we want to be able to detect future malware, especially malware variants, we also present obfuscation techniques that can be used to generate variants. These techniques can be seen on a lot of malware variants on the field. We also use the same techniques to produce our own brand of variants for our purpose. Such techniques are presented later in this chapter.

Our collection of malware samples originated from e-mails we received, malware that attacked our servers, and from various places on the internet that provide such contents. Some of the experimental data will be discussed in a later section. In order to be able to create our own variants, we must peep into how anti-virus software works: Which parts are taken into consideration when creating a signature and which bytes determine an executable is malicious. We discuss the methods used to thwart the commercial anti-virus scanners.

Figure 1. Malware taxonomy