Malware

Christian Laqua

7. Juni 2013

Inhaltsverzeichnis

1 Einleitung

Schadsoftware oder eng. Malware (das Kofferwort aus malicious (= bösartig) und Software) ist auf dem Vormarsch. Schon im Jahre 2008 wurden täglich etwa 25.000 neue Schadprogramme registriert.[1] Um dieser Situatiuon Herr zu werden, bedarf es effektiver Methoden zur automatischen Erkennung neuer Schadsoftware. Zwei Ansätze sollen im Folgenden vorgestellt werden.

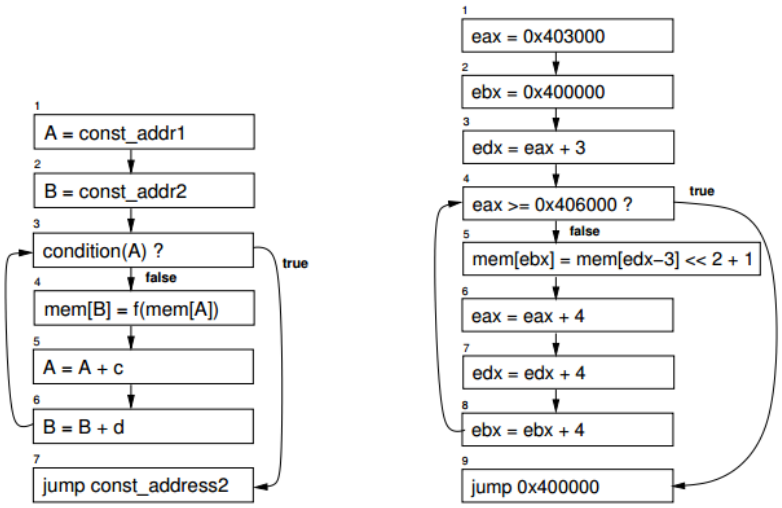
2 Mechaniken zur Malwareerkennung

Es gibt eine Vielzahl Methoden um Malware zu erkennen. Eine davon ist die statistische Analyse, welche den Binärcode untersucht. Diese Methode stößt allerdings an ihre Grenzen, sobald es darum geht, quelltextverschleierten oder sich selbst ändernden Code zu analysieren.

Eine weitere Methode ist die dynamische Malwareerkennung. Hierbei wird das Verhalten von Software zur Laufzeit beobachtet und analysiert. Kommerzielle Virenscanner setzen dabei häufig auf sogenanntes pattern-matching. Ein Programm wird hierbei als Malware erkannt, wenn eine Instruktionsabfolge mit einem regulären Ausdruck übereinstimmt. Die Schwäche hierbei ist, dass nur die Syntax analysiert wird, nicht jedoch die Semantik der Instruktionen, wodurch es für den Malwareprogrammierer oder ein Leichtes ist mit Hilfe von *Polymorphismus* und *Metamorphismus*die Syntax des Programms soweit abzuändern, dass das pattern-matching keinen Treffer mehr meldet. Ein polymorpher Virus kann seinen Schadcode verschlüsseln und dann während der Ausführung entschlüsseln. Zur Verschleierung seiner Entschlüsselungsschleife stehen dem Virenentwickler diverse Methoden zur Verfügung, wie z.B. nop-insertions, Codetranpositionen, bei denen die Instruktionsabfolge geändert und mit entsprechenden jump- Befehlen modifiziert wird, sowie Registerneuzuweisungen. Metamorphe Viren hingegen verschleiern den Kompletten Code des Virus und nicht nur den schädlichen Anteil. Replizieren sie sich, ändert sie ihren Code, so dass sie schwerer aufzuspüren sind. Der Virus transponiert seinen Code, ersetzt äquivalenten Code, ändert Sprünge und führt Registerneuzuweisungen durch. Aufgrund dieser Mechanismen benötigen Virenscanner häufige Updates, da, sobald ein Schadprogramm seine Instruktionsfolge ändert, dieser nicht mehr durch die regulären Ausdrücke erkannt. [KR08, MC05]

Eine Verbesserung stellt die semantische Malwareerkennung dar. Hierbei wird das Verhalten des ausgeführten Codes analysiert. Ein Ablauftemplate

[1] http://de.wikipedia.org/wiki/Schadprogramm



(a) Template of malicious behavior.

(b) Malware instance.

const_addr1 ← 0x403000
const_addr2 ← 0x400000
condition(X) ← X≥0x406000
f(X) ← X≪2 + 1
c ← 4
d ← 4

(c) Execution context.

const_addr1 : F(0)
const_addr2 : F(0)
c : F(0)
d : F(0)
f : F(1)
condition : P(1)

(d) Symbolic constant types.

Abbildung 1: Malwareinstanz (rechts) stimmt mit dem Template (links) überein. [MC05]

eines Malewareprogramms wird mit der ausgeführten Malwareinstanz verglichen wie in Abbildung 1 skizziert ist. Stimmt die Semantik des instantiierten Schadprogramms mit dem Template überein, so wird das Programm erfolgreich als Schadcode erkannt.[MC05]

Ein weiterer Ansatz zur verbesserten Malwareerkennung liegt darin, *automatisch* neue Malware zu erkennen und zu klassifizieren. Eine Software soll also automatisch erkennen ob eine unbekannte Schadsoftware zu einer bekannten Malwarefamilie gehört oder ob sie eine neue, bisher unbekannte Familie darstellt. Die Software muss außerdem Verhaltensweisen der (bekannten) Malewaretypen kennen, die charakteristisch für diese sind und sie unterscheidbar von anderen Malwaretypen machen. [KR08]

3 Fazit

Unter den zahlreichen Ansätzen zur Erkennung von Malware beschreiben diese beiden Verfahren zwei neue Ansätze zum Erkennen von neuem Schadcode. Während erster Ansatz dazu dient, schon bekannte, aber sich durch Umschreiben tarnende Schadprogramme durch die Semantik zu erkennen zielt der zweite Ansatz darauf ab, Malware durch Verhaltensanalyse zu Klassifizieren und entweder schon bestehenden Malwarefamilien zuzuordnen oder selbstständig als neu zu erkennen. Auch wenn beide Verfahren keine 100 prozentige Trefferquote haben, so stellen sie doch einen Schritt in die richtige Richtung dar, mit dem zunehmenden Malwaraufkommen fertig zu werden.

Literatur

[KR08] Carsten Willems Patrick Düsse Pavel Laskov Konrad Rieck, Thorsten Holz. Learning and classication of malware behavior. *D. Zamboni (Ed.): DIMVA 2008, LNCS 5137*, pages 108–125, 2008.

[MC05] Sanjit A. Seshia Dawn Song Randal E. Bryant Mihai Christodorescu, Somesh Jha. Semantics-aware malware detection. *Proceedings of the 2005 IEEE Symposium on Security and Privacy*, 2005.