

Complexity Analysis of Hostile Applets

Forensics: Using Path-Oriented Metric Analysis to Unravel Hostile Applet Algorithm Patterns, Signatures, Similarities, Authors and Derivations

Abstract

This paper uses known hostile Java applets as an example baseline that could be analyzed and profiled using path analysis to better understand the algorithms, identify their patterns and leverage the analysis to identify signatures, similarities, authors and derivations.

About Hostile Applets

The tool of choice for many hackers has been mobile code. By enticing users to a malicious web page that contains a specially crafted Java applet or application a skilled hacker can wreak havoc.

On the internet there is a comprehensive collection of hostile Java applets. A hostile applet is any applet which, when downloaded, attempts to monopolize or exploit your system's resources in an inappropriate manner. Any applet which performs, or induces you to perform, an action which you would not otherwise care to perform, should be regarded as hostile.

The power and complexity of the Java language make it extremely likely that security holes will continue to appear in years to come. While the number and nature of security holes in Java might be frightening to some, they come as no surprise to computer security professionals and hackers. Any tool as complex and as powerful as Java, being a product of human ingenuity, is bound to have numerous errors in both concept and implementation.ⁱ

Applets must obey the security rules that are put into place by the developer, using the Java security model. This security model is not infallible implementing the security model is complicated and rules are not always followed. Additionally, Java security can be sidestepped and there is always the



danger the user's computer, or company's network, can be hijacked.ⁱⁱ The most serious malicious applets find a way to circumvent Java's security mechanisms and gain complete control of the host machine. These attack applets depend on exploiting a vulnerability in a Java implementation. Other classes of

malicious applets may disturb the victim without circumventing Java's security mechanisms by behaving in an annoying or disruptive way that is within the behaviors permitted by the policy.

Java has always shown promise by being available anywhere, anytime, on any device and is nowadays also integrating with new software inventions such as web services, XML1, mobile and embedded applications. The amount of code developed using this multi-platform technology is huge and demands high security requirements. Various tools and techniques exist that can help to produce robust, reliable and secure software, however there are probably just as many methods for breaking in to code.

Why Hostile Applet Path Complexity Metrics Analysis?

Design, code, and most recently, requirements metrics have been successfully used for predicting faultprone modules.ⁱⁱⁱThey can also be used to profile and uncover security exploitable bugs.

The Cyclomatic complexity of the control flow of the program can show the characteristic style of a programmer and may suggest the manner in which the code was written. Programmers tend to show repeating patterns in their programs. It is possible to identify ownership of a program by examining source code metrics. Programmers are skillful with a limited set of constructs, mainly those that are well known to them and that allow them to write programs faster and more reliably. It is unrealistic to assume that any programmer can develop programs efficiently and correctly using an unfamiliar programming style. This does not only apply to the structure of the programs, but also to the look and feel of it.^{iv} Like naturally-evolved human languages, programming languages allow developers to express certain constructs and ideas in different ways. The differences in the way developers express their ideas can be captured in their programming styles, which in turn can be used for author identification. In the context of programming languages, it has been shown that capturing style in source code can help in determining authorship.^v

McCabe metrics have been used for years in tools used for software forensics. One of the earliest set of techniques for plagiarism detection in software is the attribute counting techniques which count the level of a certain attribute contained within a piece of code. These systems use a number of metrics such as Halstead's software science metrics and McCabe's Cyclomatic complexity.^{vi}

Being cognizant of control and data flow paths within your codebase is crucial to uncover software security vulnerabilities located off the beaten path. Security breaches are often a result of multiple interactions within the software that, on the surface, appear innocent. Criminal attackers can disrupt a system by exercising a specific sequence of interdependent decisions that produces unforeseen and possibly disastrous consequences. Analyzing control flow paths and subtree structures is crucial for both testers and developers to verify control flow integrity and uncover serious security flaws hiding in the code. This analysis may also be used to find patterns, signatures and derivations of exploits. As part of a secure, trustworthy software development process, identifying and exercising paths through the code to ensure that program behavior is correct and expected can help find the nastiest of exploits. Cyclomatic complexity and basis path analysis can be very useful in scrutinizing risky code structures and control and data flows. It can also be used to develop profiles of known exploits.

Cyclomatic Complexity has been mentioned as a possible detection method for particularly nasty bugs. In the paper "The Little Hybrid Web Worm that Could," Billy Hoffman, Lead Researcher at SPI Dynamics, and John Terrill, Co-founder of Enterprise Technology, had this to say: "One possible detection method is to examine the Cyclomatic Complexity or McCabe Complexity of a piece of arbitrary JavaScript code. The overall complexity diagram and number of closed loops should remain almost identical regardless of the number of mutations performed on the code. This follows since our mutations change the syntax of the code but not the underlying functionality then the complexity of that functionality should remain the same." The authors are investigating whether a complexity diagram alone is capable of uniquely identifying web malware."^{vii} If Cyclomatic complexity can apply to self-modifying script mutations, it can also be used for manual mutations coded by a hostile programmer, in any language. Control flow structure can be used to characterize functionality regardless of syntax variations.

Many tools used to scan for security vulnerabilities operate based on the source code syntax which can be tweaked to avert detection. Finding exploit derivations is not always easy with these tools. As others have pointed out "When the tool failed to detect a bug, it was for one of two reasons: the absence of security rules specifying the vulnerable function, or the presence of a bug in the static analysis tool. Complex code is more likely to contain complicated code constructs and obscure format string functions, resulting in lower detection rates."^{viiii} McCabe metrics have also been used to estimate the degree of protection achieved by a suite of software protection techniques^{ix}

Sometimes source code from your codebase itself is tweaked by changing, or swapping out existing code into a devious exploitable pattern. Once an Applet is written, it can be reused over and over again. Also, for novice programmers, 'slick' features can be added quite easily using code written by someone else.

Most static analysis tools only offer textual information leaving the security analyst the remaining tasks of understanding and patching the code. Visual information is believed to be helpful when fixing security flaws and using dependence graphs for pattern matching security properties has also been suggested.^x

For these reasons leveraging McCabe path-oriented analysis can help unravel hostile applet algorithmic patterns, signatures, similarities, authors and derivations.

Examples of Hostile Applet Path Complexity Metrics Analysis

Here are some of the types of hostile applets analyzed for software complexity and used as an example baseline for metric pattern match analysis.

DoMyWork.java

This Java applet makes you try to factor a moderately long integer by trial division, and it reports the results back to its home. Clearly the same could be done for many, many other sorts of calculations. While it performs no hostile actions per se, it does put your workstation to work for somebody else, perhaps a business competitor or someone trying to crack codes. To create an applet that does other sorts of work, you can replace the class GetFactor with another working class and adjust the classes Report and ReportServerSocket accordingly.

Ungreatful.java

This Java Applet tries to convince you that your system is having a security problem and that you must now login. If you do so, your user name and password are sent by the browser to the home of this applet. In any event, the applet then proceeds to the applet proceeds with a denial-of-service attack against you.

Other hostile applets include:

- * Hijacker.java (A Java trojan horse that can hijack your compiler)
- * PublicEnemy.java (A Java trojan horse that directly hacks bytecode)
- * Attacker.java (Attacks Beginner.class and makes it deviant)
- * HoseMocha.java (A Java application that defends your classes from Mocha)
- * Mutator.java (A Java application that mutates and deletes itself after 5 runs)
- * Mutator1.java (Mutator.java updated for Java Version 1.1)

Following is a report of path metrics for the source code for some of these applets.

Page 1	01/26/09	Hosti.	le Appi	let Patł	n Metrio	:5
Program	HostileAppletPat Module Name	hAnaly:	sis v(G)	ev(G)	iv(G)	gdv(G)
Page 1 Program: PublicEl Reports: Reports: Reports: Reports: Mutator ScapeGos Duple.ma: GetFacto OutPane: DoubleT TripleT Forger.: NoisyBes Hostil2 SilentFi ScapeGos Ungrates Vasteful Vasteful NoisyBes Consume DoubleT TripleT ScapeGos Ungrates ScapeGos Ungrates Consume ScapeGos Ungrates ScapeGos Ungrates Consume ScapeGos Ungrates Consume ScapeGos Ungrates Consume ScapeGos Ungrates NoisyBes Consume Consume Consume ScapeGos Ungrates NoisyBes NoisyBes NoisyBes Consume C	01/26/09 HostileAppletPat: Module Name memy.main(java.lamg: rverSocket.main(java rverSocket.main(java rverSocket.main(java rverSocket.main(java.lamg.strini frreads.run() main(java.lang.Strini fr.GetFactor() L.action(java.lang.Strini fr.GetFactor() L.action(java.lang.Strini fread.run() run() cumal(java.lang.Strini fread.sinit() cumal(java.lang.Strini freat.run() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.cun() ful.start() ful.start() ful.start() fooble.start() stort() stort() read.start() fooble.start() fooble.start() fooble.start() fooble.start() foodble.init() ar.stop() freads.stor() freads.stop() ar.update(java.awt.Gr. paint(java.awt	Hosti. hAnaly: g.Str Strin va.la ava.1 tring g[]) Event tring tring tring tring a.lan awt.E .awt. a.awt. cGrap Graph wt.Gr wt.Gr ava.1 anava.1	le App: sis ⊽(G) 2051111109988877776666665555554444444333333332222222222222	<pre>let Pat} ev(G) 817771111113113113113113111111111111111</pre>	n Metric 1♥(G) 14999977746663415334233222432222222222222222222222222	gdv(G) 3277995443276423233314233322222222222222222222222222
TripleT TripleT TripleF ErrorMes DoubleF Ungrate Ungrate DoubleT DoMyWorl Wastefu Wastefu	<pre>mreat.update(java.s rreat.paint(java.a rame.TripleFrame(j) ssage.stop() came.DoubleFrame(j) ful.stop() couble.paint(java.a ful.update(java.awt ful.paint(java.awt touble.update(java t.update(java.awt.) l.stop() couble.stop() l.update(java.awt.)</pre>	awt.G wt.Gr ava.l ava.l awt.G t.Gra .Grap .awt. Graph Graph	111111111111111111111111111111111111111	111111111111111111111111111111111111111	111111111111111111111111111111111111111	11101011111001
Wastefu AttackFi	l.paint(java.awt.G came.AttackFrame(j	raphi ava.l	1 1	1	1	0 1
Total: Average:	:		293 3.08	$137 \\ 1.44$	213 2.24	182 1.92

5

McCabe Software, Inc.

An interesting example of a distinct path complexity pattern is PublicEnemy.java. Given a target directory, this Java application searches it and all of its subdirectories for Java class files. Once a class file is located, PublicEnemy alters the contents of its access_flags for the class, its fields, and its methods. The results are the following:

- 1. The class becomes public.
- 2. Any final fields and methods become non-final; any non-public fields and methods become public; and all public fields and methods remain so. The following is the code's global data control flowgraph.



Hijacker.java is a Java Trojan that subverts Sun's javac by adding a hostile main class to the user's CLASSPATH ahead of classes.zip. In this case the subverted compiler simply announces its presence and appends the string "Hijacked!" to class files that it produces, but it could just as easily infect them with a Java Platform virus. Again this is another example of a distinct control & data flow pattern. Following is the code's flowgraph.

Graph/Listing for 'Hijacker.main(java.lang.String[])'				
Zoom In Zoom Out Print Save As Magnification Level: 2 Page 1 of 4 Hijacker.main(java.lang.String[))	Save Text	Close	Help Graph (45)%:	
Program: HostileAppletPathAnalysix0/26/09 Hijacker.mainlang.String[]) (ARMperimposed Cyclomatic Graph Unward Flows Cyclomatic 5 Essential 1 Plain Edges Design 2	Program File Language Module Letter	: HostileAppl : C:\Document : cw_Java_ins Module Name	Annotated Sour etPathAnalysis s and Settings\Tom Mo t	cce Listing
	AQ 27 28 29 30 31 32 33 34 35 36 37 38 39	Hijacker.mai AQO AQI AQ2 AQ3	<pre>n(java.lang.String[])</pre>	5 1 void main (String[] arg cay containing a hacked I
<	<		· · ·	>

1. McCabe IQ Source Code Metric Comparisons

Known Hostile Java Applets can be used as a baseline to be profiled. After the analysis is done and saved using the McCabe IQ toolset, it is fairly trivial to search within your source code to find code patterns and potential exploitable logic tweaks. Finding exact matches of patterns can be harder with lower complexity values. In addition to using McCabe path-oriented diagnostics such as Cyclomatic, essential, module design and global data complexity metrics, you may need to experiment using different elements. Path-oriented analysis and profiles can be used to identify exploits, authors and derivations. The analysis and profiles may also be leveraged to determine where in your codebase source code and logic can potentially be tweaked to create an exploit.

Specify Search Criteria	×
Title:	
Author:	
Available Elements:	Selected Elements:
call_pair ▲ call_pair_cov ▲ called_module_names ✓ called_module_types ✓ cd ✓ cdc ✓ changed_numeric ✓	ac (50) actual_parameter_names (50) actual_parameter_types (50) all_code (50) all_lines (50) blanks (50) branch (50)
Weight: S0 blanks (Blank Lines of Code)	
Save Close Help Added Metric: blanksks	

```
Date: 11/03/09
                      Module Comparison Tool Report
Program: HostileApplets2
Search Criteria: Control Structure Comparison
Search Domain:
                                     Title
        Program Name Location Date
        ----- -----
                                     ----
        HostileApplet (pcf)
Maximum similar modules: 10
Minimum similarity score: 75%
Minimum v(G) for search: 3
Total matches found: 2
Module
                                Similar Module(s)
                                                                   Similarity Score
DoMyWork.init()
                               (no matches in threshold)
DoMyWork.paint(java.awt.Graphics) (no matches in threshold)
                               Ungrateful.run() (pcf HostileApplets2) 100%
DoMyWork.run()
                      (no matches in threshold)
(no matches in threshold)
DoMyWork.start()
DoMyWork.stop()
DoMyWork.update(java.awt.Graphics) (no matches in threshold)
Ungrateful.init()
                               (no matches in threshold)
Ungrateful.paint(java.awt.Graphics) (no matches in threshold)
                               DoMyWork.run() (pcf HostileApplets2) 100%
Ungrateful.run()
                  (no matches in threshold)
(no matches in threshold)
Ungrateful.start()
Ungrateful.stop()
Ungrateful.update(java.awt.Graphics) (no matches in threshold)
```

2. Finding Control and Data Flow Similarities

A good example that can be used to illustrate a derivation can be found within the Hostile Applet baseline. Two close matches within the Hostile Java Applet codebase itself exist. Class metrics for the DoMyWork and Ungrateful Applets look similar and warrant further analysis.

Class Metric Reports for DoMyWork and Ungrateful.Applets:

Class Summary for DoMyWork:	
Number of Children:	1
Depth in Class Hierarchy:	2
Response for Class:	6
Weighted Methods Per Class:	6
Coupling of Class:	0
Number of Parents:	2
Percentage of Public/Protected Data:	100
Lack of Cohesion of Methods:	78
Accesses to Public Data:	18
Depends Upon Child:	No
Sum v(G): 13	
Avg v(G): 2.17	
Max v(G): 4	
Max ev(G): 1	
Class Member List:	
DoMyWork.paint(java.awt.Graphics)	
DoMyWork.update(java.awt.Graphics)	
DoMyWork.run()	
DoMvWork.stop()	
DoMvWork.start()	
DoMyWork.init()	
Number of non-library modules: 6	
Has Specification/Notes? No	

Class Summary for Ungrateful:	
Number of Children:	1
Depth in Class Hierarchy:	2
Response for Class:	6
Weighted Methods Per Class:	6
Coupling of Class:	0
Number of Parents:	2
Percentage of Public/Protected Data:	100
Lack of Cohesion of Methods:	76
Accesses to Public Data:	15
Depends Upon Child:	No
Sum v(G): 12	
Avg v(G): 2.00	
Max v(G): 4	
Max ev(G): 1	
Class Member List:	
Ungrateful.paint(java.awt.Graphics)	
Ungrateful.undate(java.awt.Granhics)	
Ungrateful.run()	
Ungrateful.ston()	
Ungrateful.start()	
Ungrateful.init()	
Number of non-library modules: 6	
Has Specification (Notes? No	

Here are the flowgraphs and comparison report for the Ungreatful.run module within the Ungrateful Applet and the DoMyWork.run module within the DoMyWork Applet. The path metrics and flowgraphs are strikingly similar. Upon further analysis, the path report helps isolate were the rogue code was tweaked to make a new exploit derivation.



+ 					
+=======					
+ FILE: C:\) McC: ppl:	Documents and Settings\Tom abe\Desktop\Fuzz\HostileA~ ets\Ungrateful.java	 FILE: 	C:\Documents and Settings\Tom McCabe\Desktop\Fuzz\HostileA~ pplets\DoMyWork.java		
MODULE:Ung: ========	rateful.run()	MODULE:	:DoMyWork.run()		
Ungrateful	.run() 00	Dl DoM	yWork.run()		
i 70: jav: 	a.awt.Component.repaint()	85: 	java.awt.Component.repaint() 		
73: java	a.lang.Thread.sleep(long)	88:	java.lang.Thread.sleep(long)		
73: #EX DLE]	CEPTION_THROWN# ==> UNHAN~ D_EXCEPTION	88: 	#EXCEPTION_THROWN# ==> UNHAN~ DLED_EXCEPTION		
' 82: <re<sup>- </re<sup>	turn>	, 97:	<return></return>		
* MATCH *		I			
=====================================	.run() 00	02 DoMy	yWork.run()		
70: jav: 	a.awt.Component.repaint()	85: 	java.awt.Component.repaint() 		
73: java	a.lang.Thread.sleep(long)	88: 	java.lang.Thread.sleep(long)		
' 73:#EX land	CEPTION_THROWN# ==> java.~ g.InterruptedException	, 88: 	#EXCEPTION_THROWN# ==> java.~ lang.InterruptedException		
76: con	troller == null ==> FALSE	, 91: 	controller == null ==> FALSE		
82: <re< td=""><td>turn></td><td> 97: </td><td><return></return></td></re<>	turn>	97: 	<return></return>		
* MATCH *					
Ungrateful	.run() 00	D3 DoM	yWork.run()		
70: jav	a.awt.Component.repaint()	85:	java.awt.Component.repaint()		
73: jav	a.lang.Thread.sleep(long)	88:	java.lang.Thread.sleep(long)		
73: #EX	CEPTION_THROWN# ==> NONE	 88:	#EXCEPTION_THROWN# ==> NONE		
76: con	troller == null ==> FALSE	, 91: 	controller == null ==> FALSE		
82: <re< td=""><td>turn></td><td> 97: </td><td><return></return></td></re<>	turn>	97: 	<return></return>		
* MATCH *		 =======			
Ungrateful	.run() 00	04 DoM	yWork.run()		
70: jav:	a.awt.Component.repaint()	85 :	java.awt.Component.repaint()		
73: jav	a.lang.Thread.sleep(long)	88: 	java.lang.Thread.sleep(long)		
73: #EX lan	CEPTION_THROWN# ==> java.~ g.InterruptedException	88: 	#EXCEPTION_THROWN# ==> java.~ lang.InterruptedException		
76: con	troller == null ==> TRUE	, 91:	controller == null ==> TRUE		
77: Err	orMessage	92: 	Calculator		
*** MISMAT	CH ***	 			
+==========					
ANALYSIS					
 	Number of P Number Number	oasis pasis pation	aths: 4 ches: 3		
Number of mismatches: 1 I					
Options:] +	NO DESIGN REDUCTION				

12

Examples of Hostile Java Applet Algorithm Path Signatures

Following are flowgraphs depicting the path signatures for a number of hostile Java applets.



Zoom In Zoom Out Print Save As Save Text Close Help Magnification Leve: 2 Page 1 of 4 TripleThreat.run() Program: HostileRppletPathRanalysit0/26/09 TripleThreat.run() (CR) Cyclomatic Graph Cyclomatic Graph Cyc	Graph (45)%:	Fuzz\Hosti:
Program: HostileAppletPathRnalysit0/26/09 TripleThreat.run() (CR) Superimposed Upward Flows Cyclomatic 6 Essential 4 Plain Edges Program : HostileAppletPathA File : C:\Documents and S Language: cw_Java_inst Module Module Letter Mame	Annotated Source Listing Analysis Settings\Tom McCabe\Desktop\:	Fuzz\Hosti:
	V(G)	ev(G) iv(G)
CA TripleThreat.run() 81 CAO CA1 82 83 // 84 CA2 CA3 CA4* CA5 85 85 87 CA6 CA7 CA8 CA9 CA1 88 CA15 CA16 CA17 89 90 // 91 CA20 CA21 CA22* CA2	<pre>6 public void run() { Let the applet tell its lie repaint(); Let the applet appear hones 10* CA11 CA12 CA13 CA14 CA18 try (Thread.sleep(delay catch (InterruptedExcep) Start the senseless noise 23 annov.loop();</pre>	4 t by havin(CA19);) tion e) ()





Graph/Listing for 'Wasteful.fibonacci(long)'				
Zoom In Zoom Out Print Save As Magnification Level: 2 Page 1 of 4 Wasteful.fibonacci(long)	Save Text	Close Help	Graph (45)%:	>
Program: NostileAppletPathAnalysit0/26/09 Wasteful.fibonacci(long) (CQ) Superimposed Cyclomatic Graph Upward Flows Cyclomatic 3 0 Plain Edges Essential 3 0 Plain Edges Design 21	Program File Language Module Letter	: HostileAppletPat : C:\Documents and :: cw_Java_inst Module Name	Annotated Source Listing hAnalysis Settings\Tom McCabe\Desktop\Fu v(G) ex	Azz\Hosti: 7(G) iv(G;
3 8 $9*$ 10 11 12 $13*$ 14 15 17	CQ 72 73 74 75 76 77	Wasteful.fibonacc CQD CQ1 CQ2 CQ3 CQ4 CQ5 CQ6 CQ7 CQ8 CQ9* CQ10 CQ17	<pre>i(long) 3 public long fibonacci(long k) if (k == 0 k == 1) return k; else CQ11 CQ12 CQ13* CQ14 CQ15 CQ16 return fibonacci(k - }</pre>	3 (1) + fib(
3	<			>















ⁱ Java Insecurity, Mark D. LaDue, 1996

ⁱⁱ Plain English: Risks of Java Applets And Microsoft ActiveX Controls, Jennifer M. Marek GIAC Security Essentials Certification (VER 1.3), 4 March 2002, SANS Institute

iii "Comparing Design and Code Metrics for Software Quality Prediction" Yue Jiang, Bojan Cukic, Tim Menzies, Nick Bartlow, The Lane Department of Computer Science and Electrical Engineering, WVU

iv "Authorship Analysis: Identifying The Author of a Program" Ivan Krsul Eugene H. Spafford, The COAST Project - Department of Computer Sciences, Purdue University

v Identifying Source Code Authorship, Robert Lange, Jonathan Max-Sohmer, Maxim Shevertalov, Jay Kothari, Spiros Mancoridis, Software Engineering Research Group, Department of Computer Science

vi "SOURCE CODE AUTHORSHIP ANALYSIS FOR SUPPORTING THE CYBERCRIME INVESTIGATION PROCESS" Georgia Frantzeskou, Stefanos Gritzalis Laboratory of Information and Communication Systems Security, Aegean University, Stephen G. MacDonell, School of Computer and Information Sciences Auckland University of Technology

vii "The Little Hybrid Web Worm that Could" Billy Hoffman, SPI Dynamics, John Terrill Co-founder Enterprise Management Technology

viii "Measuring the Effect of Code Complexity on Static Analysis Results" James Walden, Adam Messer and Alex Kuhl, Department of Computer Science Northern Kentucky University

ix "Realistic and Affordable Cyberware Opponents for the Information Warfare BattleSpace" Martin R. Stytz, Ph.D., Sheila B. Banks, Ph.D., Michael J. Young, Ph.D., Air Force Research Laboratory Wright-Patterson AFB, OH 45431

x "Pattern Matching Security Properties of Code using Dependence Graphs" John Wilander and Pia Fak, Dept. of Computer and Information Science, Linkopings Universitet