

Web Security Research at Indiana University

Dr. XiaoFeng Wang Associate Professor School of Informatics and Computing Indiana University



Our Adventures on the Web



- Privacy: get your health records, salary, investment secret from Web apps' side channel
 - Microsoft Buddies: Shuo Chen and Rui Wang
- Flaws: shop for free, log into your web account through 3rd-party Web APIs
 - Microsoft Buddies: Shuo Chen and Rui Wang
- Misdeeds: how to advertise infections and click frauds through the New York Times
 - Microsoft Buddies: Yinglian Xie and Fang Yu



Knowing Your Enemy: Understanding and Detecting Malicious Web Advertising

Zhou Li, Kehuan Zhang, Yinglian Xie, Fang Yu and XiaoFeng Wang



JOBS REAL ESTATE AUTOS ALL CLASSIFIEDS WORLD U.S. POLITICS NEW YORK BUSINESS DEALBOOK TECHNOLOGY SPORTS SCIENCE HEALTH OPINION ARTS Books Movies Music Television Theater STYLE Dining & Wine

Fashing & Chile

Scramble as **Credit Tightens** Across Europe

By ERIC DASH and NELSON D. SCHWARTZ As European banks pull

back on lending, companies around the globe are finding it harder to borrow, edging the world economy toward another slump.

- · Obama Meets Leaders of the European Union
- Dire Warnings Are Building on European Debt Crisis

CAMPAIGN 2012

Barney Frank, a Top Liberal, Won't Seek Re-election

By ABBY GOODNOUGH Mr. Frank, who first



Ionica Lopossay for The New York Times

Black Workers Struggle in Public Sector

By TIMOTHY WILLIAMS

Since the recession's declared end, middle-class blacks like Pamela Sparks, a postal worker in Baltimore, above, have seen a traditionally prosperous job source run dry.

DEALBOOK

Money Found in Britain May Belong to MF Global

What Happened on the Border?

A transparent investigation of the NATO strikes, with Pakistan's participation, is essential.

- · Brooks: The Life Reports II
- · Bruni: Silvio's Postscript
- · Nocera: Germany Cuts Off Its Nose
- Cohen: Doctrine of Silence
- Fish: Dogs and Cars
- Op-Ed: Optimism in Egypt
- · Room for Debate: Legal Rights for the Obese?

MARKET	rs » At	2:04 AM E
JAPAN		CHINA
Nikkei	HangSeng	Shangha
8,477.82	18,302.43	2,411.0
+190.33	+264.62	+27.9
+2.30%	+1.47%	+1.179
Data d	elayed at leas	t 15 minute



Gift Subscriptions The New York Times is offering digital gift subscriptions to NYTimes.com and tablet and smartphone apps. Subscriptions are available for 12 weeks or 26 weeks and start at \$30.

· Order Now » Frequently Asked Questions



Today Only!



[http://www.adexchanger.com/pdf/Display-Advertising-Technology-Landscape-2010-05-03.pdf]

Sersonal Personal	Antivirus				
	Personal	rus		Get fu	II time protection
		Scanning for threats			
OV DV	erview				
Q Vin	us scan	Full computer	scan	Remo	ve threats
Lice	ense				
	date product	File Name	Re	sult/Infection	Â
HQ		C:\Boot\memtest.exe	In	fected: I-Worm.Sober.i - Tro	ijan 🔤
		C:\config.sys	Int	fected: Suspicious.Harakit - 1	Trojan, Virus
		C: Vhiberfil.sys	Inf	fected: Suspicious.Harakit + 1	Trojan, Virus
		C:\pagefile.sys	Int	fected: Suspicious.Harakit - 1	Trojan, Virus 👻
		* Lin	111		
Spyware DB Version: Status:	8: 8531 8.2.10.25 Not activated	Current object: C:\Program F	iles (Lenovo (System L	Jpdate\session\7bg675ww	Stop
5 L ESTATE OS	Credit Tightens		- AND	What Happened on the Border? A transparent investigation	
CLASSIFIEDS	Across Europe		10 A	of the NATO strikes with	
RLD	By ERIC DASH and NELSON D.			Pakistan's participation is	Cift Subscriptions
ia Mi	SCHWARTZ			essential	The New York Times is
ITICS	As European banks pull		200	costruit.	offering digital gift
N YORK	back on lending, companies			Brooks: The Life Reports II	subscriptions to
INFSS	around the globe are		A	 Bruni: Silvio's Postscript 	NVTimes com and table
LBOOK	finding it harder to borrow,			 Nocera: Germany Cuts Off 	and emartahone anne
HNOLOGY	edging the world economy			Its Nose	Subscriptions are avail
DTS	toward another slump.			Cohen: Doctrine of Silence	Subscriptions are availa
ENCE	· Obama Meets Leaders of the		SN /	 Fish: Dogs and Cars 	IOF 12 Weeks OF 26 Week
ALTU	European Union			• Op-Ed: Optimism in Egypt	and start at \$30.
NION	Dire Warnings Are Building on	Monica Lopossa	ay for The New York Times	· Room for Debate: Legal	Order Now »
NION	European Debt Crisis	Black Workers Struggle in Pu	blic Sector	Rights for the Obese?	 Frequently Asked Question
1S		By TIMOTHY WILLIAMS			
ian	CAMPAIGN 2012	Since the recession's declared end, mid	dle-class blacks like	MARKETS » At 2:04 AM ET	
ic		Pamela Sparks, a postal worker in Balt	imore, above, have	JAPAN CHINA	SAVE 50
vision	Barney Frank, a Top	seen a traditionally prosperous job sour	rce run dry.	Nikkei HangSeng Shanghai	es Times Digital Subserie
ater	Liberal, Won't Seek			8,477.82 18,302.43 2,411.01	on a rimes Digital Subscrip
YLE	Re-election	DEALBOOK		+190.33 +264.62 +27.98	Today Only!
ing & Wine	By ABBY GOODNOUGH	Money Found in Britain May Belor	ng to MF Global	+2.30% +1.47% +1.17%	totaly only i
shine 9 Shile	Mr. Frank, who first			Data delayed at least 15 minutes	



Malvertising via Nytimes





Defense: What has been done

- Ad behavior restriction [AdSafe, Finifter'10, Louw'10]
 - Limited applicability to different forms of attacks: drive-bydownload, phishing, click-fraud
- URL features and domain reputations [Zhang'11, John'11]
 URLs can be easily modified by attackers
 Domains can be hijacked
- Code analysis [Cova'10]
 - Obfuscate code to evade detection
 - Leverage ad syndication to bypass the code checking



What hasn't been done

Understanding:

≻ How serious is the problem?

> What does malcontent delivery path look like?

➢ Roles of ad nodes? Topologies?

Detection:

Complement existing techniques with infrastructure information?



The First Step We Made

- Measurement of Malvertising
 - > Malvertising: drive-by download, Scam, Click fraud
 - Scale, node/path features of malcontent delivery infrastructures
- Detection of Malvertising
 - Path segment based detection
 - It works: caught 15 times more cases than popular blacklist/malware scanners (Safe Browsing, Forefront)

Get down to the specifics: Data Collection

- From June 21st to September 30th
- 12 Virtual Machines with instrumented browser
- Alexa top 90,000 web sites visited regularly
- Extract ad redirection paths

Indiana University School of

nformatics

Node: freeonlinegames.com doubleclick.net/abc, advertising.com/abc



Path: freeonlinegames.com -> doubleclick.net/abc -> adsloader.com/abc

Case: freeonlinegames.com -> doubleclick.net -> adsloader.com



Some Statistics

- 24 million ad paths
- 22 million nodes, >90% ad nodes
- Scanned with Forefront and Google Safebrowsing
- 543 malicious nodes, 263 domains
- 938 malicious cases
- 286 infected publishers (Ranked from 314 to 89184)
- Long-lived campaign (2 months), short-lived domain (3 days)

Example: a Fake AV Campaign





Node Features

- Most of malicious nodes have unknown roles
 - >90% malicious nodes, <8% legitimate nodes
- Registered within a year, expire in a year
 - >70% malicious domains, <20% legitimate domains
- Free domain providers like .co.cc used widely
- Follow URL patterns
 - /showthread.php\.php $?t=d{8}$ matches 34 domains



Properties of Malicious Pairs

Malicious node pairs appear less frequently



Properties of Malicious Paths

- Longer path length (8.11 > 3.59 of legitimate paths)
- Ad syndication is the major problem (>60% cases)
- The closer to bad nodes, the more suspicious

Indiana University School of

informatics





Our Ideas for Detection

- Analyze ad-delivery sequences
 Focus on short subsequences
 Annotate nodes with rich attributes
- Use statistical learning to generate detection rules
 > Adapt to new, ever changing attacker strategies



Detection Framework



Results

Testing June - September

phishing pages	56	0.00%		_
drive-by-download pages	172	9.88%	• 5% FD	R
click-fraud pages	155	10.97%	• 15x ne	w findings
all pages	326	8.90%	• 10.5 d	avs early detection
phishing cases	104	0.00%	than ac	of obrowsing
drive-by-download cases	1171	6.23%	lian se	alebiowsing
click-fraud cases	4221	4.10%		
all cases	5496	4.48%		
		Sa	afebrowsing	&
			Forefront	
	#MadTracer	#S&F	%FP	%New Findings
phishing pages	#MadTracer 12	#S&F 0	%FP 9	%New Findings 100.00%
phishing pages drive-by-download pages	#MadTracer 12 216	#S&F 0 104	%FP 9.00% 0.00% 9.26%	%New Findings 100.00% 51.85%
phishing pages drive-by-download pages click-fraud pages	#MadTracer 12 216 89	#S&F 0 104 7	%FP 9.00% 0.00% 9.26% 14.61%	%New Findings 100.00% 51.85% 92.13%
phishing pages drive-by-download pages click-fraud pages all pages	#MadTracer 12 216 89 291	#S&F 0 104 7 111	%FP 9.26% 9.26% 14.61% 11.00%	%New Findings 100.00% 51.85% 92.13% 61.86%
phishing pages drive-by-download pages click-fraud pages all pages phishing cases	#MadTracer 12 216 89 291 23	#S&F 0 104 7 111 0	%FP 9.26% 9.26% 14.61% 11.00% 0.00%	%New Findings 100.00% 51.85% 92.13% 61.86% 100.00%
phishing pages drive-by-download pages click-fraud pages all pages phishing cases drive-by-download cases	#MadTracer 12 216 89 291 23 627	#S&F 0 104 7 111 0 216	%FP 9.26% 9.26% 14.61% 11.00% 0.00% 13.88%	%New Findings 100.00% 51.85% 92.13% 61.86% 100.00% 65.55%
phishing pages drive-by-download pages click-fraud pages all pages phishing cases drive-by-download cases click-fraud cases	#MadTracer 12 216 89 291 23 627 3422	#S&F 0 104 7 111 0 216 42	%FP 0.00% 9.26% 14.61% 11.00% 0.00% 13.88% 3.65%	%New Findings 100.00% 51.85% 92.13% 61.86% 100.00% 65.55% 98.77%

Testing October

New Click-Fraud: Hijack User Traffic





Conclusion

- Malvertising is a big issue
 - \geq 1% top publishers are infected
- Study on infrastructures can lead to a promising new direction
 - > 15x more coverage, 5% false positive
 - Discover new attacks
- Usage and deployment
 - Ad exchange service (e.g., Ad-center): capture malicious and fraudulent ad entities
 - > Anti-virus (e.g., ForeFront): provide new malware signatures
 - > End users (you and me): detect and stop ongoing exploits



