Measuring the Role of Automation in Malicious Web Activities

Thesis presentation of Xigao Li

Committee:

Michael Ferdman
Aruna Balasubramanian
Alexandros Kapravelos
Nick Nikiforakis (Advisor)
Amir Rahmati (Advisor)

Thesis Statement

“Attackers have extensively employed automation techniques to conduct malicious web activities.

Thus, it is imperative for defenders to employ automation techniques in order to detect, understand, and mitigate the impact of such activities.”

What is malicious web activity?

“Interactions with web servers and web users that result in *negative* impacts.”

Examples

1. Brute-force log-in attempts

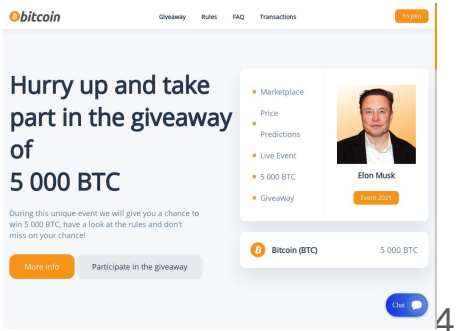
- Try to log-in with a list of usernames and passwords
- Compromise user security
- Incur web server overhead

2. Scanning for web server vulnerabilities

- Reconnaissance/attack with a list of known vulnerabilities
- Compromise web server security

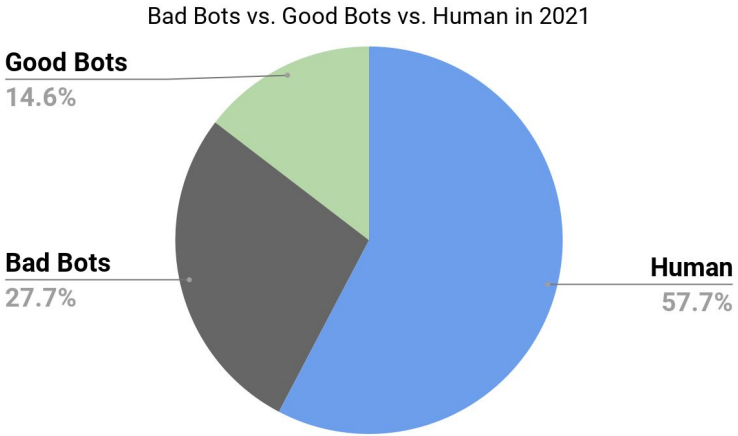
3. Conduct scam activities

- Defraud users for funds with cryptocurrency giveaway scams
- Entice user to invest with automated comment posts

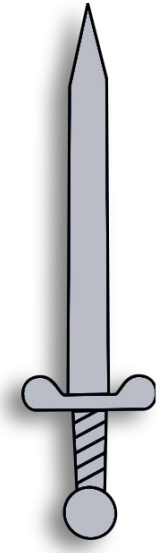


The role of Automation …

- Internet activities can be (in fact, most of them are) automated.
- Programs that run those automated tasks are referred as **web bots**.



Industry report: 65% bots are malicious



Data Source: 2022 Imperva Bad Bot Report: Evasive Bots Drive Online Fraud.
https://www.imperva.com/resources/resource-library/reports/bad-bot-report/

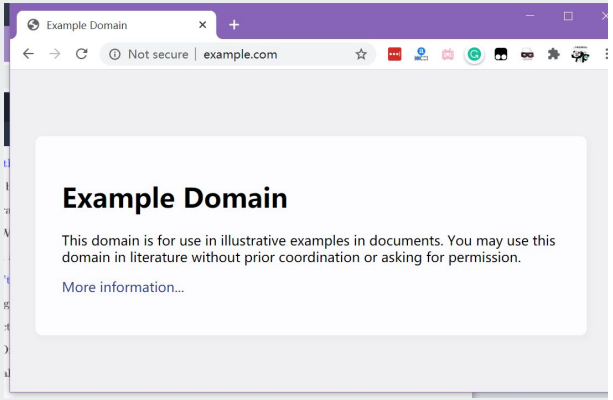
Automation for good

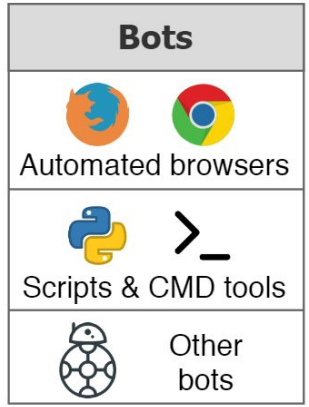
Defenses can also be automated.

- Build clean, large datasets - measure/understand malicious activities
- Build detection systems - detect/prevent malicious activities

Different types of web bots

```
import requests
requests.get("https://example.com")
```





A simple one-liner script

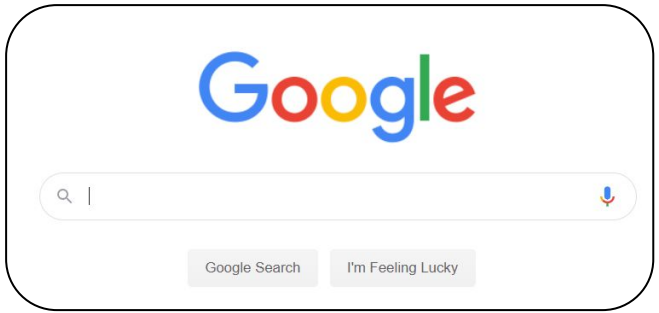
Complicated browser

- Controlled by program
- Perform clicks,take screenshot

Types of benign bots

Benign bots contribute to the Internet.

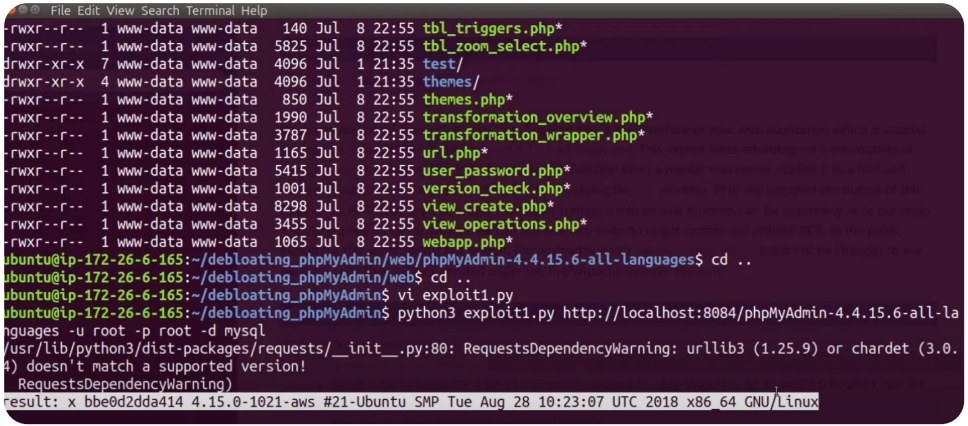
- Content indexing - Googlebot, Bingbot …
  - Compute/provide ranking (e.g. Alexa)
  - Content analysis
  - Power products
- Academic Research
- Cache/Rehost service
  - Internet Archive



Types of malicious bots

Malicious bots cause damage to servers and users

- Credential stuffing attacks
- Probing for vulnerabilities
  - Fingerprint application
  - Steal unprotected information
  - Exploit discovered vulnerabilities
- Denial-of-Service attacks
  - Impact a website's availability



Example of exploiting CVE-2016-5734 through web requests (arbitrary code execution)

Bots identifying as...

Not all bots identify themselves honestly.

- Spoofing User-Agents
- Browsing with automated browsers
- Use proxy to change IP address

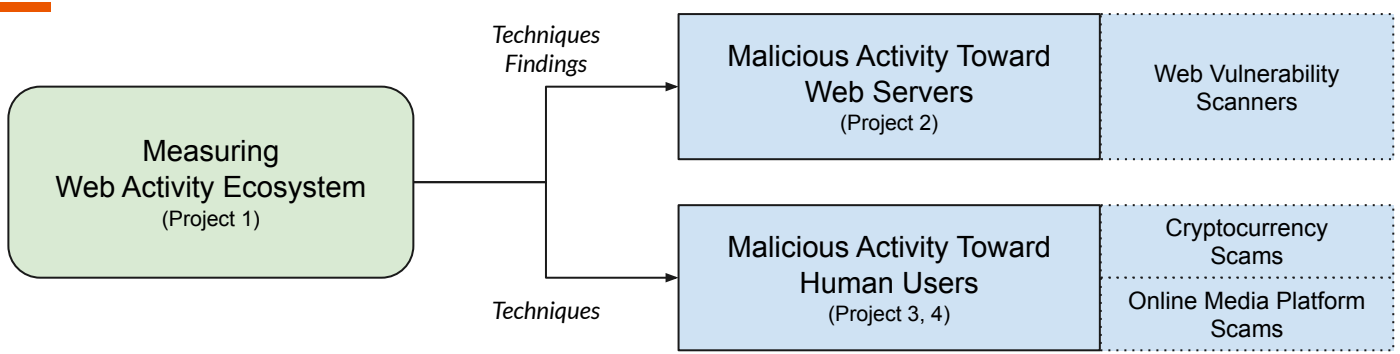
Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_3) AppleWebKit/537.75.14 (KHTML, like Gecko) Version/7.0.3 Safari/7046A194A

Mozilla/5.0 (Linux; Android 4.4.2; Nexus 4 Build/KOT49H) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/34.0.1847.114 Mobile Safari/537.36

Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/84.0.4147.89 Safari/537.36

User-Agent: a string field in HTTP requests, user for self-identification of the client to the server

Presentation Roadmap



1. **Characterizing Automated Browsing Activities (IEEE S&P 2021)**
2. Understanding and Detecting Unwanted Vulnerability Scanning (ACM WWW 2023)
3. Understanding and Detecting Cryptocurrency Giveaway Scams (NDSS 2023)
4. Characterizing Comment Scams on Media Platforms (NDSS 2024, in submission)

Good bot, Bad Bot:

Characterizing Automated Browsing Activities

Published at IEEE Symposium on Security and Privacy (S&P) 2021

Overview

Design and build Aristaeus *

- A system that provide flexible remote deployment and management of honeysites.
- Design high-interaction honeysites,
  - Full functional web applications, equipped with state-of-the-art fingerprint and identification techniques

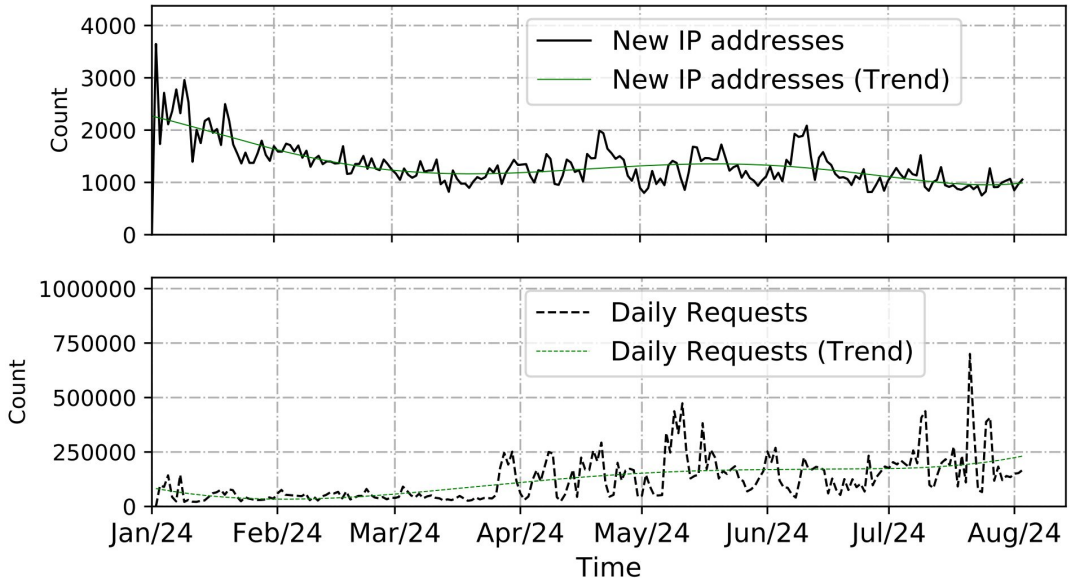
A systematic study on the internet bot traffic

- 7 months of study with 100 honeysites
- Capture, fingerprint and uncover bot activities through various traces

** Minor God in Greek mythology , creator of arts like bee keeping*

Result: Bot Traffic Analysis

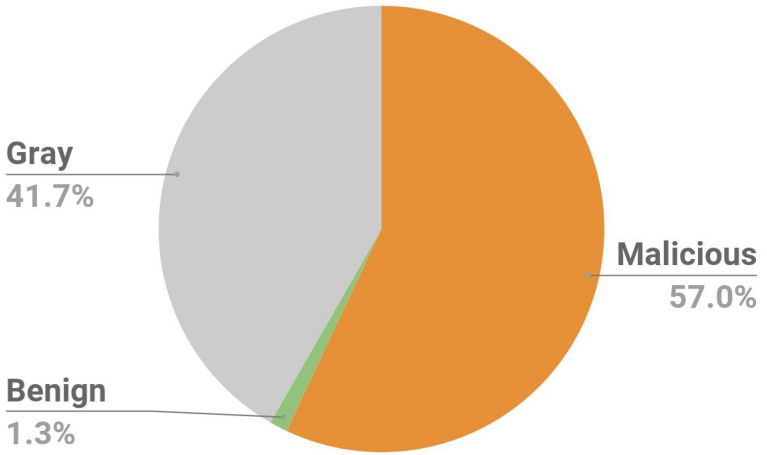
- We keep observing traffic from new IP addresses, for the entire 7 months
- Average 1,235 requests per day per honeysite



Result: Bot Intentions

Bots are categorized as “Benign”, “Malicious”, “Other/Gray”.

- **Benign**
  - Asking for valid resources similar to a normal browser
  - No intent of attacking
- **Malicious**
  - Send unsolicited POST requests towards authentication endpoints
  - Send invalid requests trying to exploit vulnerabilities.
- **Other/Gray**
  - None of the above traits



Bots are pretending to be browsers

Bots claiming to be:

- **Chrome: 82.6% are fake**
  - Mostly curl/wget
- **Firefox: 98.5% are fake**
  - 60.6% are go-http-client
  - 34% are libwww-perl
  - Remaining 5.4% are still not firefox

TLS fingerprinting is effective against evasion / cloaking.

Takeaways

- Any online website will receive 1,000+ requests/day, ~1% are benign
- 98% bots are rudimentary HTTP libraries, pretending to be browsers
- Bots prefer low-hanging fruits, aiming at easiest vulnerabilities
- Only 13% of bot IP appeared in IP blocklists
- Exploits that go public are quickly abused - Within a few hours
- TLS fingerprinting is effective against cloaking and evasion

Result: Bad bots

- Credential bruteforce attempts
  - 50.8% of total requests
  - 47,667 unique IP addresses
  - 99.6% of bots issued fewer than 10 attempts

- **Target Reconnaissance attempts**
  - Application fingerprinting
  - Exploitation attempts
  - Scanning open-access backdoors
  - Scanning for unprotected sensitive files

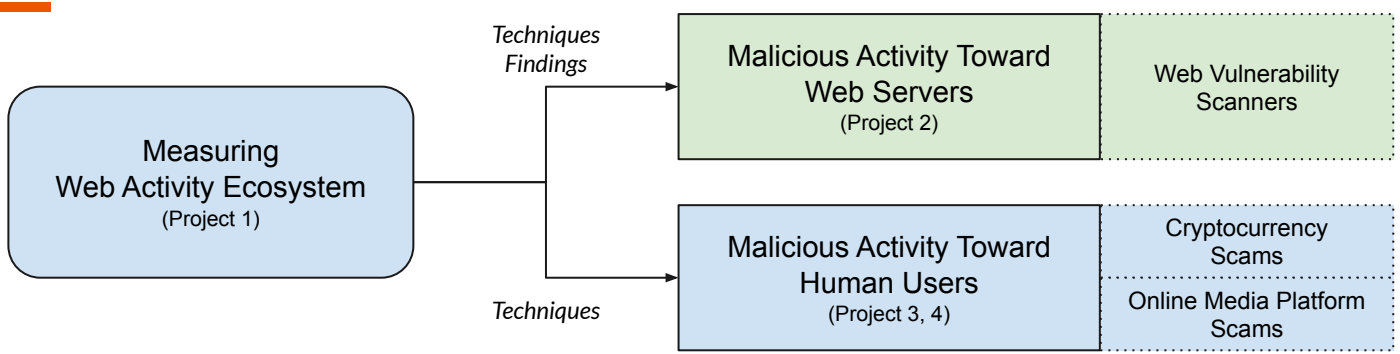
Result: Bad bots

- Credential bruteforce attempts
  - 50.8% of total requests
  - 47,667 unique IP addresses

What is the role of automation in malicious activities toward web servers?

  - Application fingerprinting
  - Exploitation attempts
  - Scanning open-access backdoors
  - Scanning for unprotected sensitive files

Presentation Roadmap



1. Characterizing Automated Browsing Activities (IEEE S&P 2021)
2. **Understanding and Detecting Unwanted Vulnerability Scanning (ACM WWW 2023)**
3. Understanding and Detecting Cryptocurrency Giveaway Scams (NDSS 2023)
4. Characterizing Comment Scams on Media Platforms (NDSS 2024, in submission)

Scan Me If You Can:

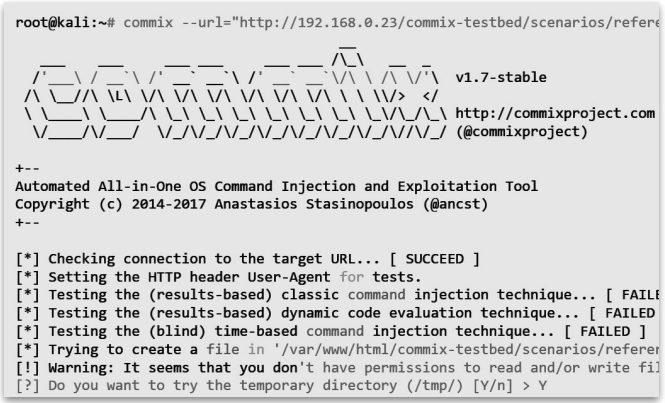
Understanding and Detecting Unwanted Vulnerability Scanning

Published at ACM TheWebConf (WWW) 2023

Web vulnerability scanner (WVS)?

Automated, “point-and-click” tools that scan web applications for vulnerabilities.

- **Perfect tool for penetration testers**
  - Identify and fix low-hanging vulnerabilities
- **Full-auto weapon for malicious actors**
  - Identify and exploit low-hanging vulnerabilities



Commix Scanner Example

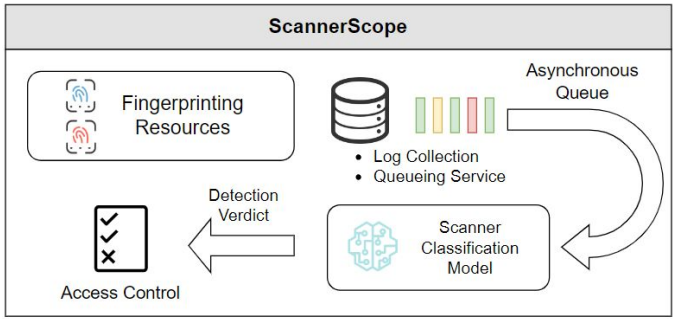
Overview

We developed a testbed for WVS:

- Automatically launching WVS and scan our own targets
- Observed differences between WVS and users though user study

We designed ScannerScope:

- Use supervised machine learning model classifies users vs. WVSs



Threat Model

Malicious actors abusing off-the-shelf WVS:

- Scan without permission of website owner

Allow web administrators to apply access-control policy:

- Block IP address
- Throw CAPTCHA

Testbed and Data Collection

WVS: 12 WVSs are evaluated.

- **Top open-source WVS of top OWASP pentesting tools and academic scanners**

Human: 159 Users are included in the test.

- **Users are asked to perform randomized tasks**

Scanner Name	Version
WPScan(kali)	3.8.13
Arachni	1.5.1
OWASP Zap	D-2020-12-21
WMap	1.5.1
Wapiti	3.0.3
Nikto	2.1.6
W3af	1.6.45
Skipfish (kali)	2.10b
Commix	2.9-stable
Google Tsunami	0.0.5
Black Widow	N/A
Enemy of the State	N/A

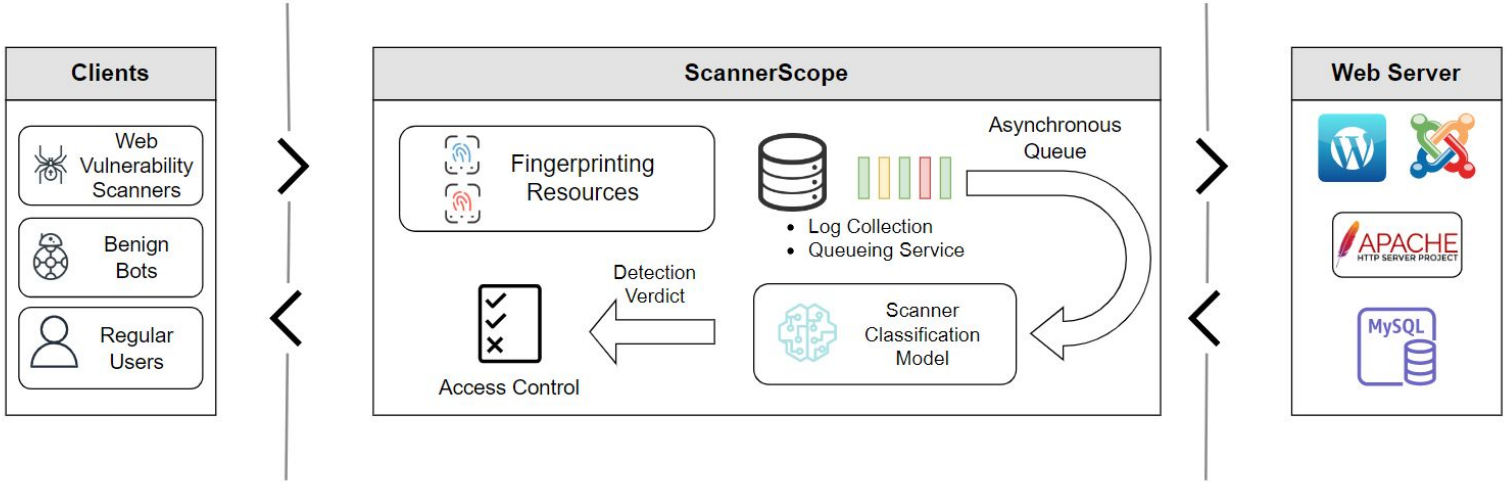
Eriksson et al., *Black widow: Blackbox data-driven web scanning*. IEEE S&P 2021
Doupé et al., *Enemy of the state: A state-aware black-box web vulnerability scanner*, Usenix Security 2012

Scanner behaviors

- The majority of scanners send a large number of requests.
- Some WVSs have distinct exploration and attack phases.
- Some WVSs only used a subset of attack vectors in each execution.
- WVSs focus on different endpoints than human users, producing a large number of invalid requests.

The observed differences between human users and WVSs inspired the design of ScannerScope.

ScannerScope Design



ScannerScope is designed as a reverse proxy.

ScannerScope achieved 99% accuracy on both WordPress and Joomla web traffic.

Takeaways

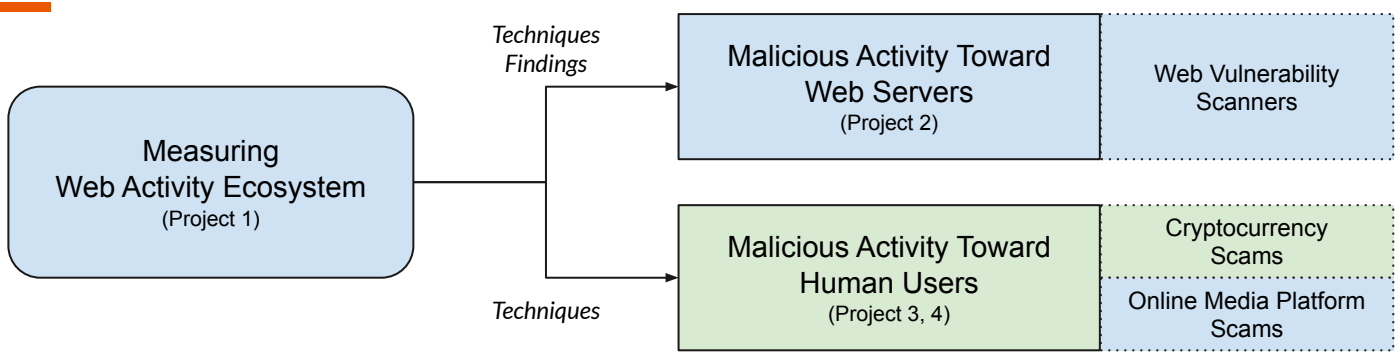
- Automation can be used by malicious actors to scan web vulnerabilities
- Automation can be used by defenders to detect those scanners.

Takeaways

- Automation can be used by malicious actors to scan web vulnerabilities
- Automation can be used by defenders to detect those scanners.

What is the role of automation in malicious activities toward *human users*?

Presentation Roadmap



1. Characterizing Automated Browsing Activities (IEEE S&P 2021)
2. Understanding and Detecting Unwanted Vulnerability Scanning (ACM WWW 2023)
3. **Understanding and Detecting Cryptocurrency Giveaway Scams (NDSS 2023)**
4. Characterizing Comment Scams on Media Platforms (NDSS 2024, in submission)

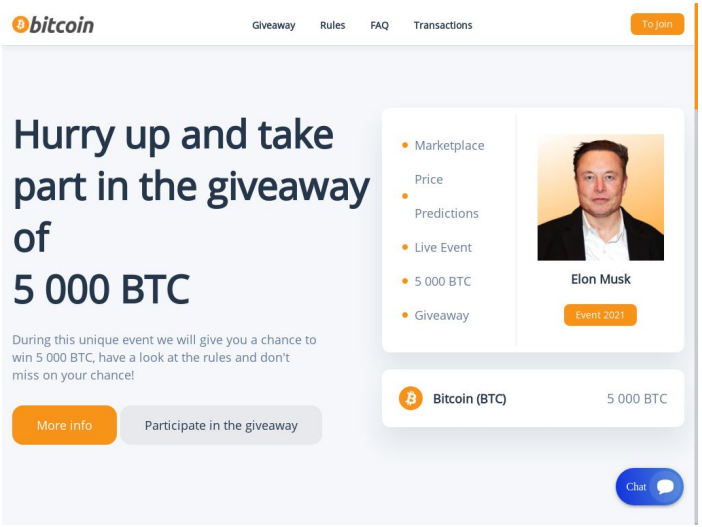
Double and Nothing:

Understanding and Detecting Cryptocurrency Giveaway Scams

Published at Network and Distributed System Security Symposium (NDSS) 2023

Introducing Cryptocurrency Giveaway Scams

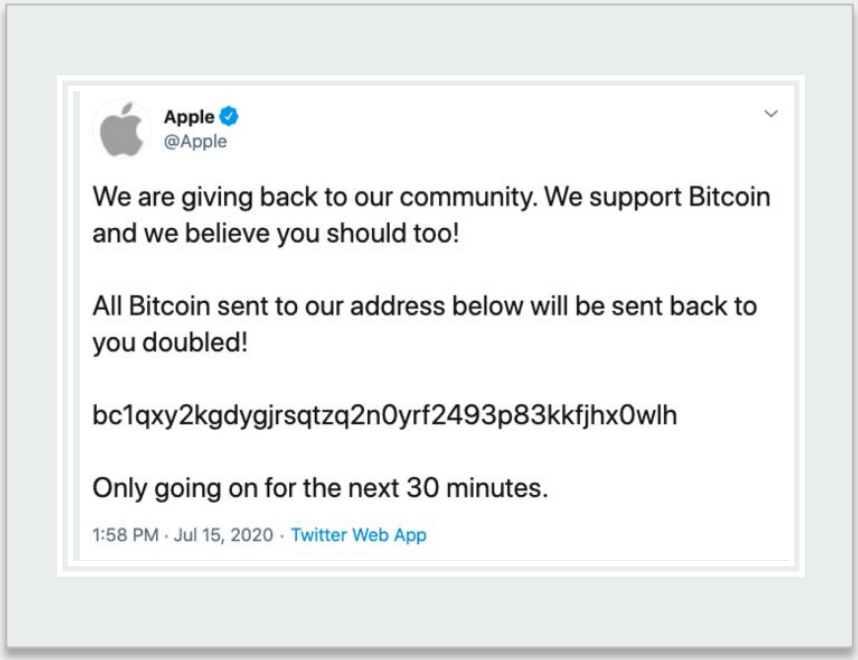
- Professional-looking websites
- Abuse names and images of celebrities
- Advertise “giveaway events” that promise to multiply user funds
- Require cryptocurrency fund transfer to a specific wallet address



Introducing Cryptocurrency Giveaway Scams

- **Advertising scams**
  Scammers advertise scams through social media accounts, and YouTube channels

- **Most famous event**
  2020 twitter hack - 130 accounts belonging to high profile individuals tweeting the scam
  - **Celebrities affected:** Barack Obama, Joe Biden, Bill Gates, Warren Buffett, Jeff Bezos, Michael Bloomberg, etc.



There are no large-scale studies of cryptocurrency giveaway scams - people solely rely on user reports or incident investigations.

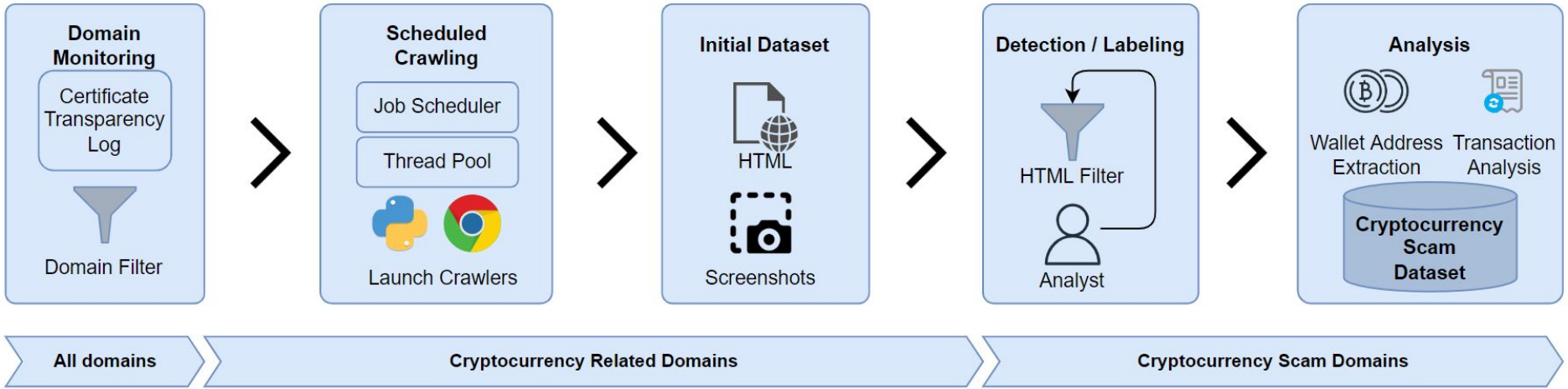
Overview

- First large-scale measurement of cryptocurrency scam websites
  - Design and build CryptoScamTracker
    - System to identify and record cryptocurrency-giveaway scams through Certificate Transparency logs
  - Captured 10,079 scam websites in 6 months

- First quantitative analysis of cryptocurrency scams
  - Tens of millions of dollars were stolen
  - Found clear signs of automation in setting up scam pages

CryptoScamTracker Design

- CryptoScamTracker is composed of 3 modules:
  - Domain monitoring module
  - Crawl and detection module
  - Analysis Module



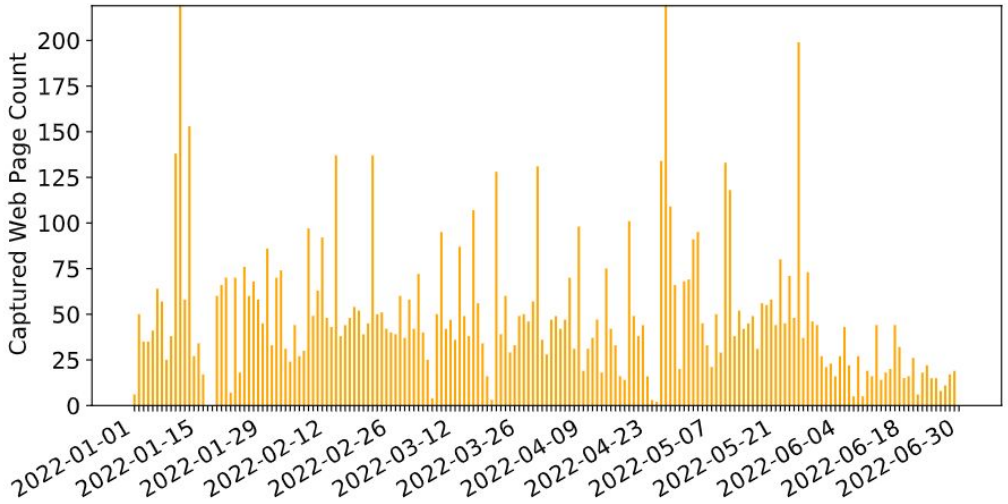
CryptoScamTracker Design

- Domain monitoring module
  - Monitor Certificate Transparency (CT) logs with a keyword filter
- Crawl and detection module
  - Issues requests for suspicious scam domains, retrieve HTML and screenshot of web pages, acquire domain information from WHOIS
  - Detect and store scam webpages by scam keyword filter and presence of cryptocurrency wallet
- Analysis module
  - Analyze HTML, images, transactions, etc.

Dataset collection

- Collected 6 months of data from January 1, 2022 to July 1, 2022.
- 10,079 cryptocurrency scam web pages
- 3,863 domains, 2,712 IP addresses
- 2,266 scammer wallet addresses extracted

Details of cryptocurrency scams



- Average of 55.7 new scam web pages each day
- No significant correlation with market price and daily captures.

Domain analysis: domain name

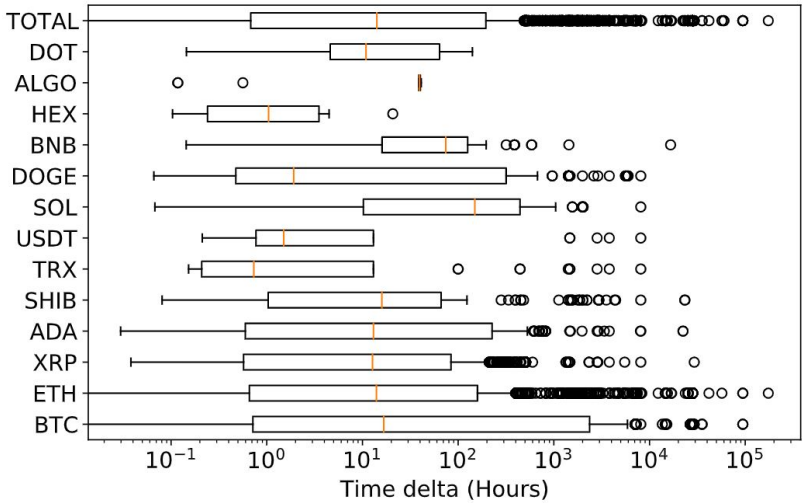
- Scam operators prefer traditional gTLD for domains
  - .com, .org, .net
  - Total registration cost: $22,000+
- Scam domains tend to use year-related keyword
  - 22-shib.com, 2022-ethereum.org
  - 38% domains contain “22” or “2022”, 0.31% contain “21” or “2021”
  - 34.89% domains contain multipliers like “2x” or “3x”

TLD	Domain Count	Total estimated cost
com	1435	10274.6
org	762	5836.92
net	618	3083.82
us	156	154.44
info	127	247.65
live	113	212.44
io	74	2126.76
online	49	48.51
gift	42	556.08
tech	36	79.2
(Total)	3412	22620.42

Domain analysis: registration info

- Names / Personal Emails are available in WHOIS info
  - Can be used for clustering scams into campaigns
- Cryptocurrency scam websites prefer non-popular hosting providers
  - Reg.ru, DDoS-Guard, etc.
  - DDoS-GUARD hosting 9.47% of all scam websites

    yet only 0.05% of benign top 10K websites
- Most domains have short “lifespan”
  - 50% websites have lifespan less than 26 hours
  - One domain was registered at 2002, 6 years before the concept of Bitcoin

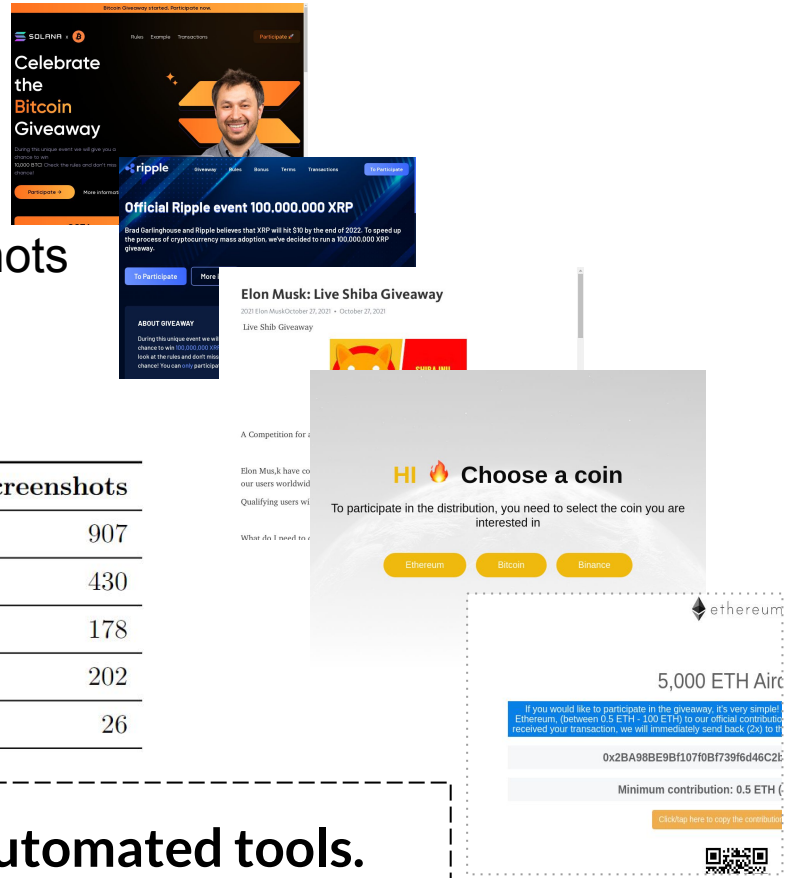


Webpage analysis: JavaScript

- Common JavaScripts are identified from scam web pages.
  - JQuery (12,795) - basic JavaScript library
  - Live chat services (8,372) - free chat-as-a-service library, which scammers used for persuading victims
  - Animation Libraries (2363) - present smooth animation
  - Analytics (399) - Google / Yandex analytic metrics
  - Website Obscurity (476) - prevents user to inspect web page source

Live chat service can become an early-warning system against scams.

Webpage Layout



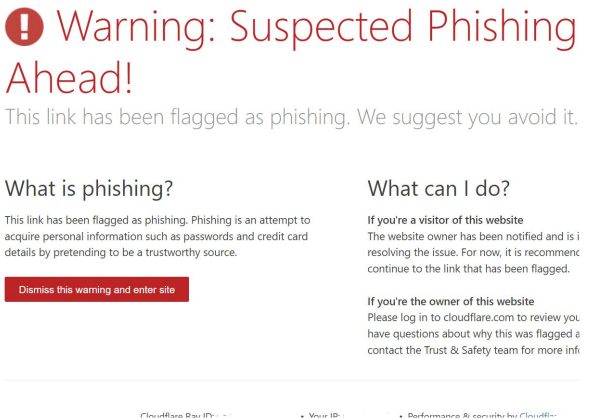
- Use perceptual hashing over 3,832 screenshots
  - Ultimately group screenshots into 139 clusters
- Image Clusters have 5 different styles

Style #	Style Detail	Clusters	Screenshots
1	Scam web page with celebrity portrait	44	907
2	Scam web page without celebrity portrait	22	430
3	Media article style	8	178
4	"Fork" style with two or more cryptocurrency	14	202
5	QR Code visible in first page style	2	26

Scam websites are created via automated tools.

Anti-scam techniques

- Online crowd-sourcing database: only captures a small percentage of domains and wallets in our dataset (CryptoScamDB: 0.35%, BitcoinAbuse: 14%)
- Domain blocklists: Only 16.75% domains we captured are marked suspicious/malicious by VirusTotal.
- Hosting provider regulations: Scammers evade regulations by using unpopular hosting providers (e.g REGRU, DDoS-Guard).



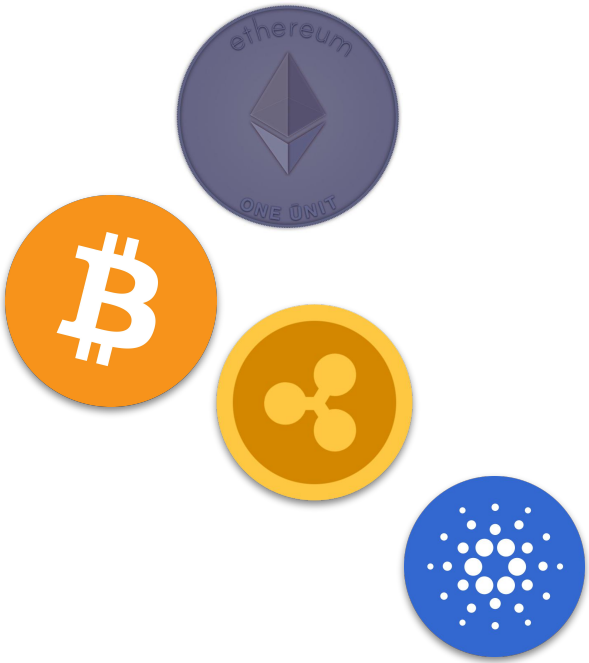
Anti-scam techniques have limited coverage.

Targeted Cryptocurrency

- Total of 13 cryptocurrencies are targeted
- Most favored cryptocurrency:
  - Ethereum (ETH) - 6,777 scams
  - Bitcoin(BTC) - 5,980 scams
  - Ripple (XRP) - 1,303 scams
  - Cardano (ADA) - 818 scams

  Top 4 cryptocurrencies attracted 90% of the scam websites in our dataset.

- Scammer may set up multiple cryptocurrency in one domain



Funds stolen (BTC, ETH, ADA, XRP)

- Scammers' wallets are publicly accessible on blockchain, allowing us to track all past transactions.
- $24.9M–$69.9M funds were stolen by Scammers
  (using the minimum and maximum cryptocurrency prices during our study)
- Total Stolen Cryptocurrency:
  - BTC: 940.07
  - ETH: 4,330.26
  - ADA: 2,141,876.52
  - XRP: 5,799,593.93

The most successful ETH scammer received a total of 258.54 Ethereum, could be worth of $990,000 in 2022.

Takeaways

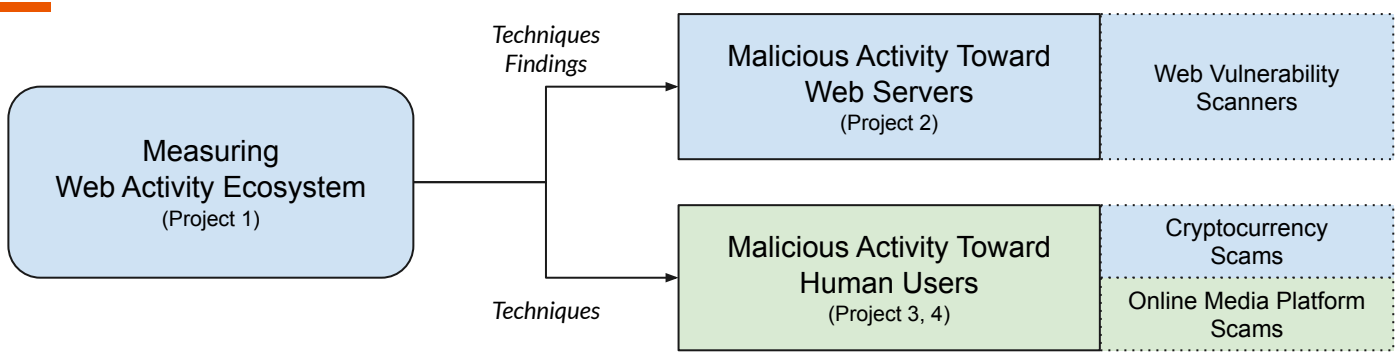
- CryptoScamTracker is effective in capturing cryptocurrency scam websites
- 10K scam web pages served from 3.8K domains are captured in our study
- $24.9M–$69.9M funds were stolen by Scammers
- Websites screenshots are similar, indicating they are built from automated tools
- Blocklists and Online DBs have limited coverage
- Third-party JavaScript libraries may be a future way detecting scams

Takeaways

- CryptoScamTracker is effective in capturing cryptocurrency scam websites
- 10K scam web pages served from 3.8K domains are captured in our study
- $2
- Websites screenshots are similar, indicating they are built from automated tools
- Blocklists and Online DBs have limited coverage
- Third-party JavaScript libraries may be a future way detecting scams

What does automation do in other scam activities that directly involved scammers?

Presentation Roadmap



1. Characterizing Automated Browsing Activities (IEEE S&P 2021)
2. Understanding and Detecting Unwanted Vulnerability Scanning (ACM WWW 2023)
3. Understanding and Detecting Cryptocurrency Giveaway Scams (NDSS 2023)
4. **Characterizing Comment Scams on Media Platforms (NDSS 2024, in submission)**

Like, Comment, Get Scammed:

Characterizing Comment Scams on Media Platforms

Submitted to Network and Distributed System Security Symposium (NDSS) 2024

Introducing Comment Scams

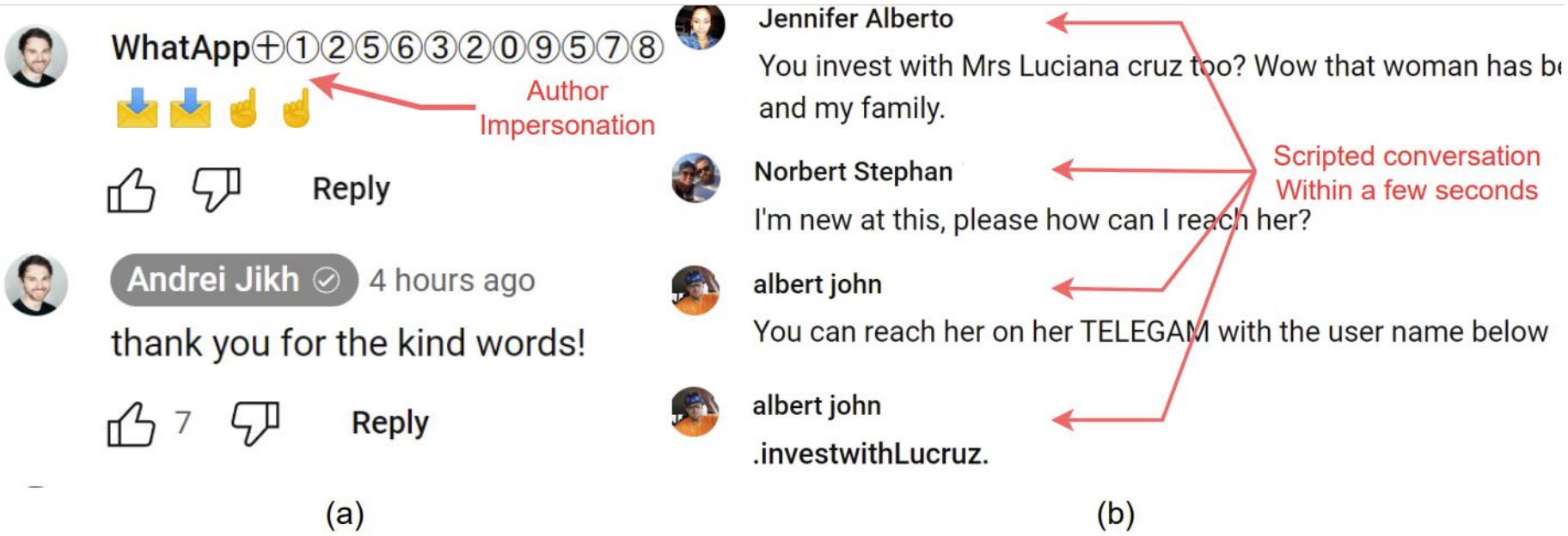
Comment scam on media platforms

- Comments or replies, enticing users to contact them through messages
- Solicit a chance to win a gift or investment opportunities

- Example: *“TextMe on WhatsApp (555)-5555”*

.

Example of Comment Scam



- **Scammers apply multiple tactics to evade platform regulation.**

Overview

- Build a reliable infrastructure monitoring YouTube comments
  - Monitor past and new videos in specific YouTube channels
  - Periodically take snapshots of comment section
- Design heuristic filters to identify scam comments
  - Text-based filters (Textual)
  - Image-based filters (Graphical)
  - Time-based filters (Temporal)

Dataset Collection

- Measurement range: October 1st, 2022 to March 31st, 2023
- Monitored Channels: 20
- Videos: 8,226
- Captured comments: 8.8 Million
- Filtered scam comments: 206K (2.34% of total comments)

Comment Scam Features

- **Textual -** Scammers use Visually Similar Symbols (VSS) to evade automated detection systems
- **Graphical -** Scammers apply similar profile images to impersonate channel owners
- **Temporal -** Scammers split the conversation and even contact phone numbers, and use multiple accounts to post them together to form a fabricated short story

Comment Scam Features

MESSAGE ME ON TELEGRAM +1234 ✓ (ASCII latin)

MESSAGE ME ON TELEGRAM ᐩ1234 ✗ (Latin Letter Small Capital Unicode)

whatsapp 1234 ?

whatsapp 1234 ?

Visually Similar Symbols (VSS)

- a **(U+0061) vs a (U+1D5BA)**

Comment Scam Features

MESSAGE ME ON TELEGRAM +1234	✓	(ASCII latin)
MESSAGE ME ON TELEGRAM ⁺¹²³⁴	✗	(Latin Letter Small Capital Unicode)
whatsapp 1234	?	(ASCII latin letters)
whatsapp 1234	?	(Mathematical Sans-Serif Small Unicode)

Visually Similar Symbols (VSS)

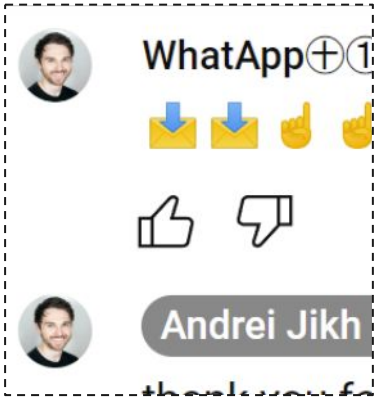
- a **(U+0061) vs a (U+1D5BA)**
- **Most common ways scammers used to evade detections**
- **Difficult to identify by unaware users**

Comment Scam Features

- **Textual -** Scammers use Visually Similar Symbols (VSS) to evade automated detection systems
- **Graphical -** Scammers apply similar profile images to impersonate channel owners
- **Temporal -** Scammers split the conversation and even contact phone numbers, and use multiple accounts to post them together to form a fabricated short story

Comment Scam Features

- **Graphical -** Scammers apply similar profile images to impersonate channel owners
  - Difficult to distinguish in the view of inexperienced users
  - Perceptual hashing to compare with channel owners



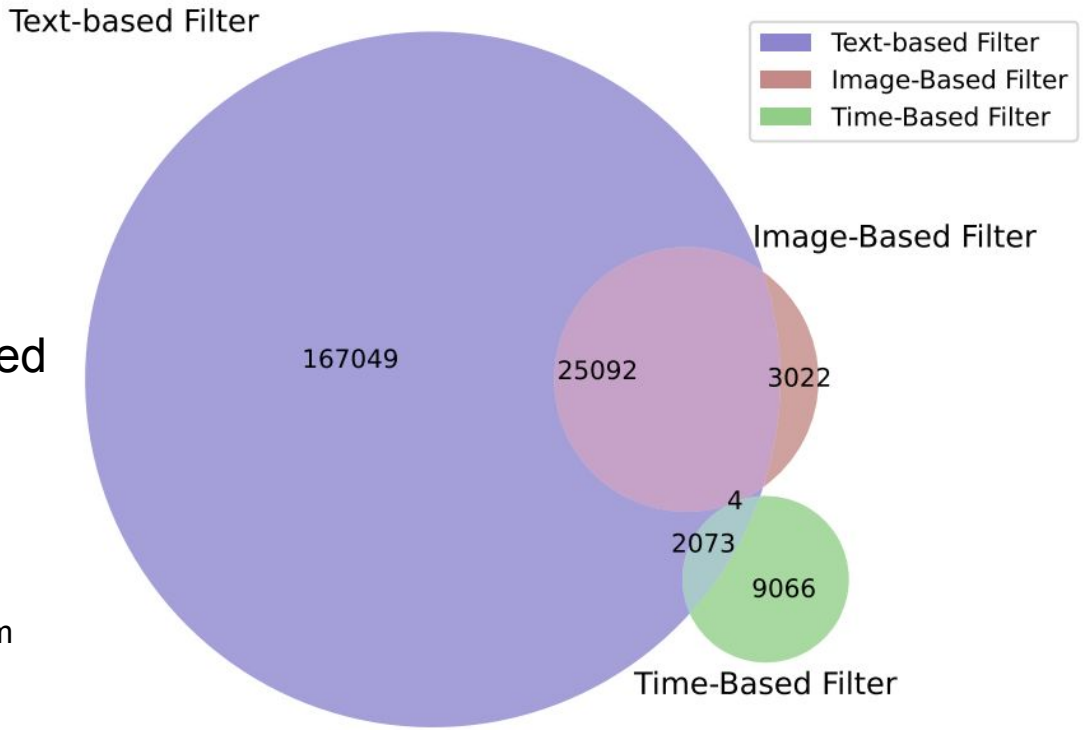
Comment Scam Features

- **Textual -** Scammers use Visually Similar Symbols (VSS) to evade automated detection systems
- **Graphical -** Scammers apply similar profile images to impersonate channel owners
- **Temporal -** Scammers split the conversation and even contact phone numbers, and use multiple accounts to post them together to form a fabricated short story

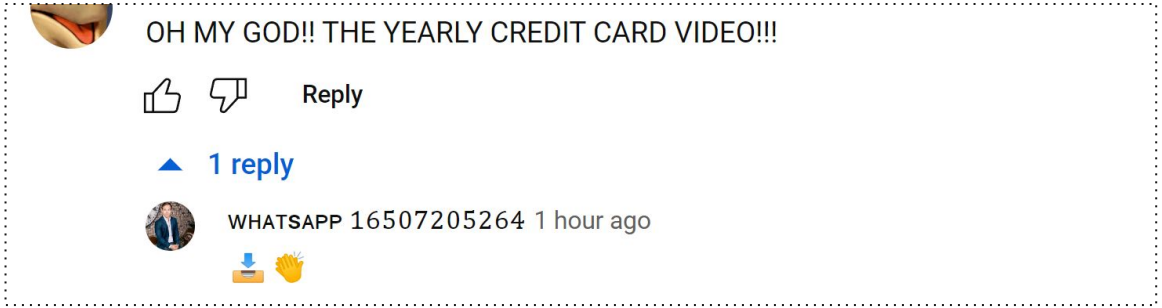
Filter results

- Text-based filters captured majority of scam comments
- A single comment can be labelled with multiple filters
- Filters have intersections

(Scammers use multiple ways to evade platform regulations)



Comment Scam Features



Scammer text

- Convey general information (no specific target)
- Entice user to contact (on other platforms)
- Impersonate or fabricate (increase credibility)
- Automated through scripts (widespread)

Scam Campaigns

Campaign ID	Accounts	Comments Posted	Affected Videos	Targeted Channels	Affected Categories
1	112	4045	92	1	Finance
2	59	703	324	4	News/Politics, Finance
3	46	5405	66	2	Finance
4	45	692	321	4	News/Politics, Finance
5	44	5662	76	2	Finance

Connect campaigns by phone numbers and account IDs

- Largest campaign have 112 accounts
- Most widespread campaign targeted 324 videos
- Only 31.42% scam accounts were deactivated during study

Interacting with scammers

- IRB-approved study
- Pretend to be unaware victims and send text message to 50 scammers
- Explore scammer tactics and payment channels
- Platform: WhatsApp and Telegram

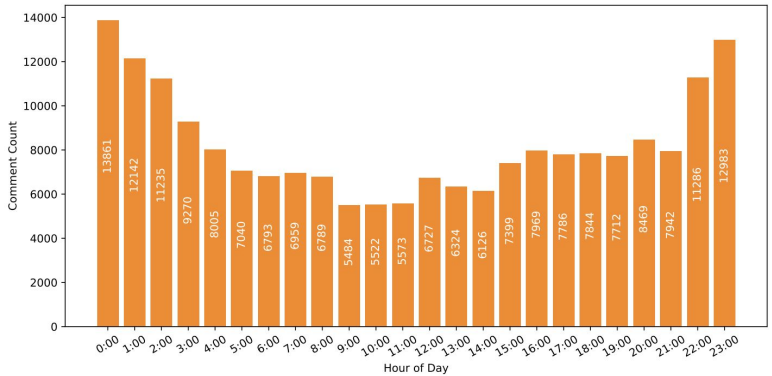
Scammer tactics / payment channels

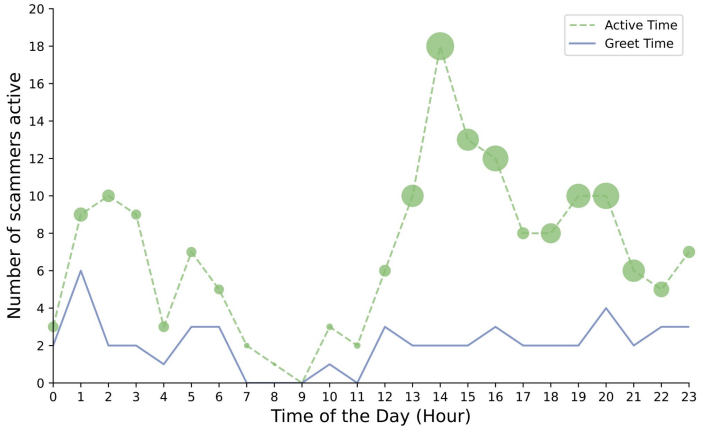
- **Cryptocurrency Investment (76%)**
  - Promise unrealistic high-yield investments (15% to 1300% weekly return)
  - Impersonation as channel owner or broker
  - Entice user to transfer cryptocurrency to scammer's wallet

- **Fake Prize (22%)**
  - Promise a prize (usually related to channel content)
  - Request shipping charges ($50 to $500)

- **Others (2%)**

Scammer working time





- Scam comments are mostly published at 12AM (0:00), probably due to API quota reset
- Some scammers working in different timezone than United States despite their numbers are mostly U.S. based.

Funds stolen (cryptocurrency)

Crypto-currency	# of Wallets	Total Amount of Cryptocurrency	USD Value (Min. - Max.)
Bitcoin (BTC)	31	67.64	$1.07M - $1.92M
Ethereum (ETH)	16	36.49	$0.04M - $0.07M
(Total)	47	-	$1.11M - $1.99M

Millions of dollars (equivalent) were stolen by only 31 scammers

Automation in Defense

- Track transactions
- Textual, Graphical and Temporal Filters
- Can be used to automatically flag comments for verification

Result: Role of automation in scam activities

- Automation is widely abused by malicious actors to scam human users on Internet.
- Automation can also play a role as measurement and defense toward scam activities.

Conclusion

The role of automation are two-sided in web activity. It can be used for:

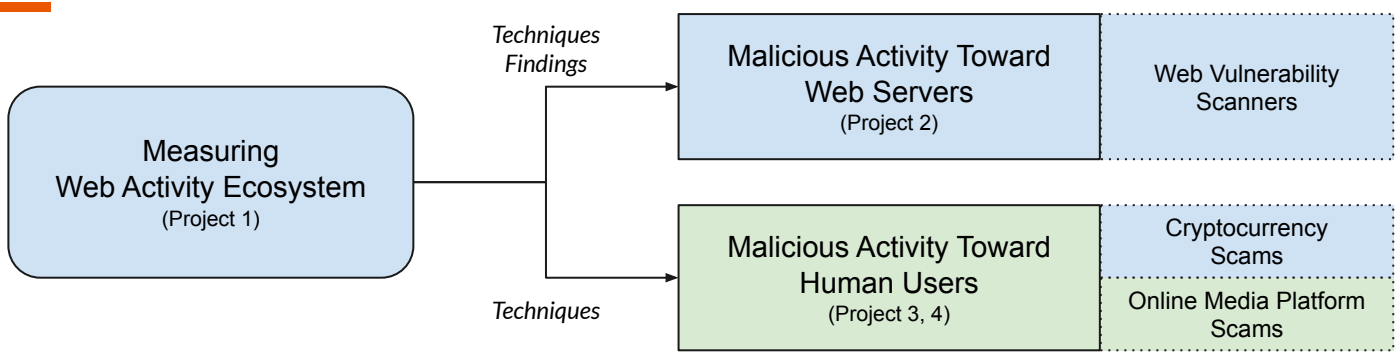
- Toward web servers (scanning, exploiting)
- Toward human users (setting up scams)

While attacker can leverage such automation techniques, defenders can also use them to:

- Understand malicious activities (measurements)
- Detect malicious activities (defending systems)

Leveraging automation techniques could lead to an upper hand in the constant arms-race.

Measuring the Role of Automation in Malicious Web Activities



1. Characterizing Automated Browsing Activities (IEEE S&P 2021)
2. Understanding and Detecting Unwanted Vulnerability Scanning (ACM WWW 2023)
3. Understanding and Detecting Cryptocurrency Giveaway Scams (NDSS 2023)
4. Characterizing Comment Scams on Media Platforms (NDSS 2024, in submission)