Fraunhofer
FKIE

Fraunhofer Institute for
Communication, Information
Processing and Ergonomics FKIE

2024-04-26 · **Steffen Enders**, **Daniel Plohmann**, Manuel Blatt

A Taxonomic Overview of Prevalent Malware Communication Strategies

Motivation

Motivation

Why is Malware C2 even that interesting?

Motivation

Why do we need a taxonomy for Malware C2?

elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys
elf.mirai	win.azorult	win.darkcomet	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws	win.pikabot	win.smokeloader
win.agent_tesla	win.brute_ratel_c4	win.dbatloader	win.lumma	win.privateloader	win.socks5_systemz
win.amadey	win.bumblebee	win.dcrat	win.masslogger	win.qakbot	win.stealc
win.asyncrat	win.cloudeye	win.emotet	win.nanocore	win.quasar_rat	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire	win.recordbreaker	win.vidar
☺		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader
		win.glupteba	win.panda_stealer	win.remcos	win.xworm

A Taxonomy for Malware C2 Strategies

Taxonomy for Malware Communication Strategies

Current State of the Art?

Taxonomy of Botnet Threats

A Trend Micro White Paper / November 2006

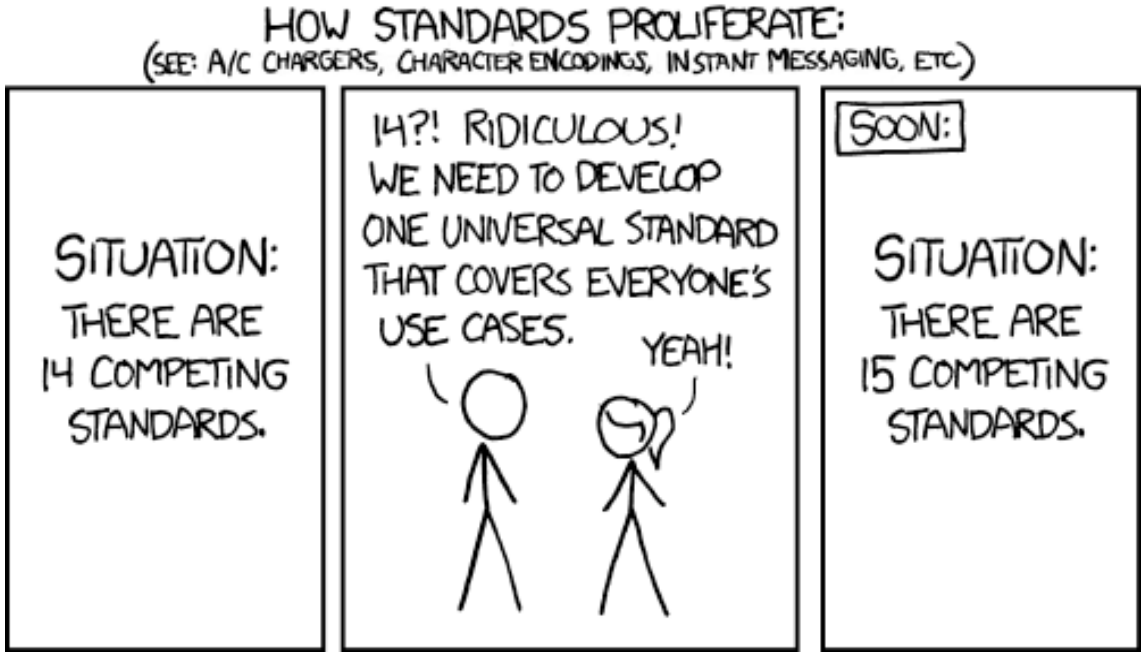
Botnets: Lifecycle and Taxonomy

Nabil Hachem*, Yosra Ben Mustapha*, Gustavo Gonzales Granadillo* and Herve Debar*
*Institut TELECOM
Telecom SudParis
CNRS Samovar UMR 5157
Evry, France
Email: {nabil.hachem, yosra.ben_mustapha, gustavo.gonzalez_granadillo, herve.debar}@it-sudparis.eu

ATT&CK®

Taxonomy for Malware Communication Strategies

Current State of the Art?



https://xkcd.com/927

Malware C2 Taxonomy

Overview

Taxonomy, loosely based on Trend Micro Report 2006:

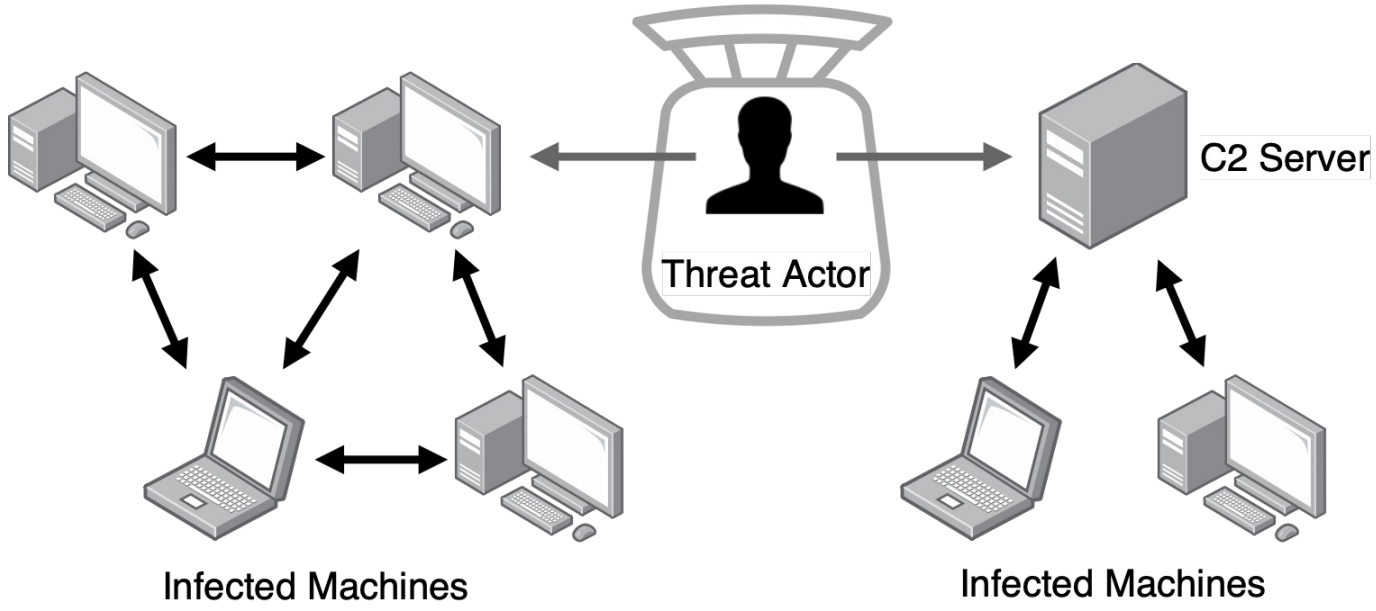
1. **C&C Models**
2. **Rally Mechanisms**
3. **Communication Behavior** ← **Added**
4. **Carrier Communication Protocols** } **Properly differentiated**
5. **C&C Protocols** } **Properly differentiated**
6. **Evasion Techniques**
7. ~~Botnet-Specific Stuff~~ ← **Removed**

Malware C2 Taxonomy

C&C Models

- **High-level classification**
- **Structure of the overlay network**

C&C Models	
Centralized	Star
Centralized	Hierarchical
P2P	
Hybrid	
Random	



Malware C2 Taxonomy

Rally Mechanisms

Rallying Mechanisms	
IP-Address	Hardcoded
IP-Address	Seeding
Domain Names	Hardcoded
Domain Names	DGA
Further Indirection	

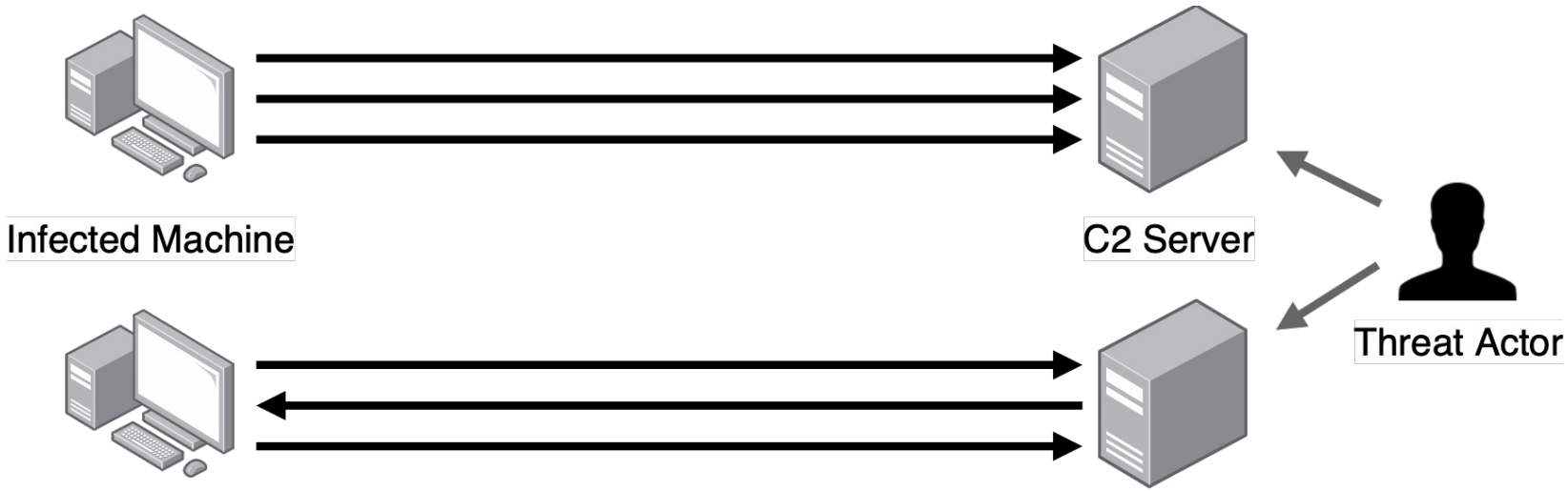
- **Initial Establishment of C2 Communication Channel, e.g. through:**
  - **Hard-coded IPs/Domains**
  - **Domain Generation Algorithms**
  - **Indirection via Web Services:**

Malware C2 Taxonomy

Communication Behavior

Communication Behavior	
Transmission	Push/Pull
Sessions	1-way/2-way

- **Direct botmaster to bot commands vs. Bots periodically checking in**
- **Bots receiving commands w/o providing responses vs. 2-way communication**

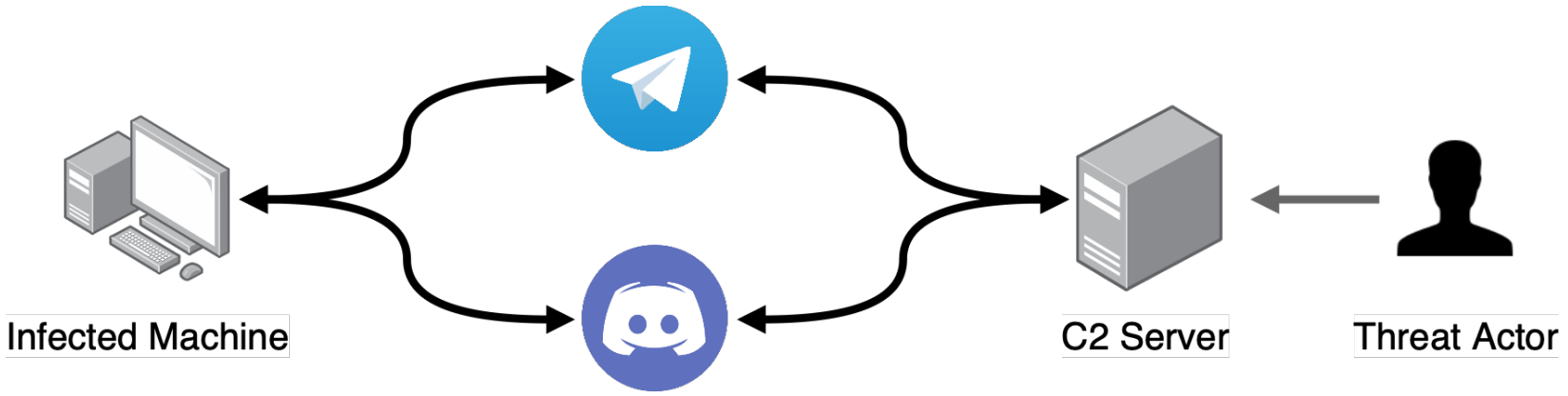


Malware C2 Taxonomy

Carrier Communication Protocols

Protocols to transmit customized data

- **Direct, low-level network communication**
- **Encrypted raw socket communication**
- **Utilization of common protocols or web-services**

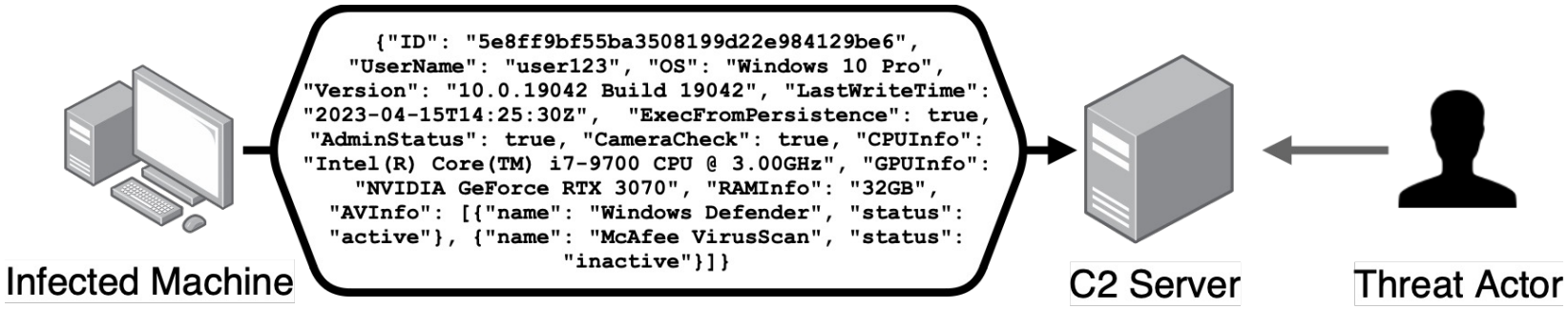


Carrier Protocols
Raw TCP/UDP Socket
Raw SSL/TSL Socket
Application Layer Procols
Web Services

Malware C2 Taxonomy

C&C Protocols

- **Protocol of the transmitted data**
- **Highly specific, not categorized any further**

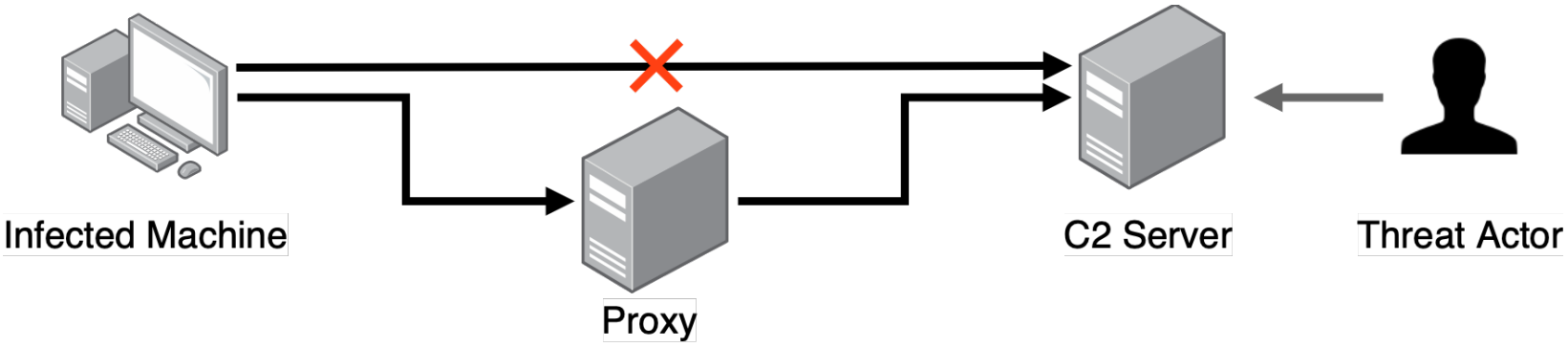


C&C Protocols
Custom Text-Based
Custom Binary
Existing Protocols
...

Malware C2 Taxonomy

Evasion Techniques

- **Backup C&C Channels.**
- **Masking traffic origin**
- **Rapidly changing proxies**
- **Concealing C&C data**



Evasion Techniques
Fallback Channel
Proxy Servers
Stepping Stones
Fast-Flux
DNS Calculation
Encryption
Data-Obfuscation
Non-Standard Ports
Traffic Pattern Manipulation

Malware C2 Taxonomy

Conclusion

C&C Models	
Centralized	Star
Centralized	Hierarchical
P2P	
Hybrid	
Random	

Rallying Mechanisms	
IP-Address	Hardcoded
IP-Address	Seeding
Domain Names	Hardcoded
Domain Names	DGA
Further Indirection	

Communication Behavior	
Transmission	Push/Pull
Sessions	1-way/2-way

C&C Protocols
Custom Text-Based
Custom Binary
Existing Protocols
...

Carrier Protocols
Raw TCP/UDP Socket
Raw SSL/TSL Socket
Application Layer Procols
Web Services

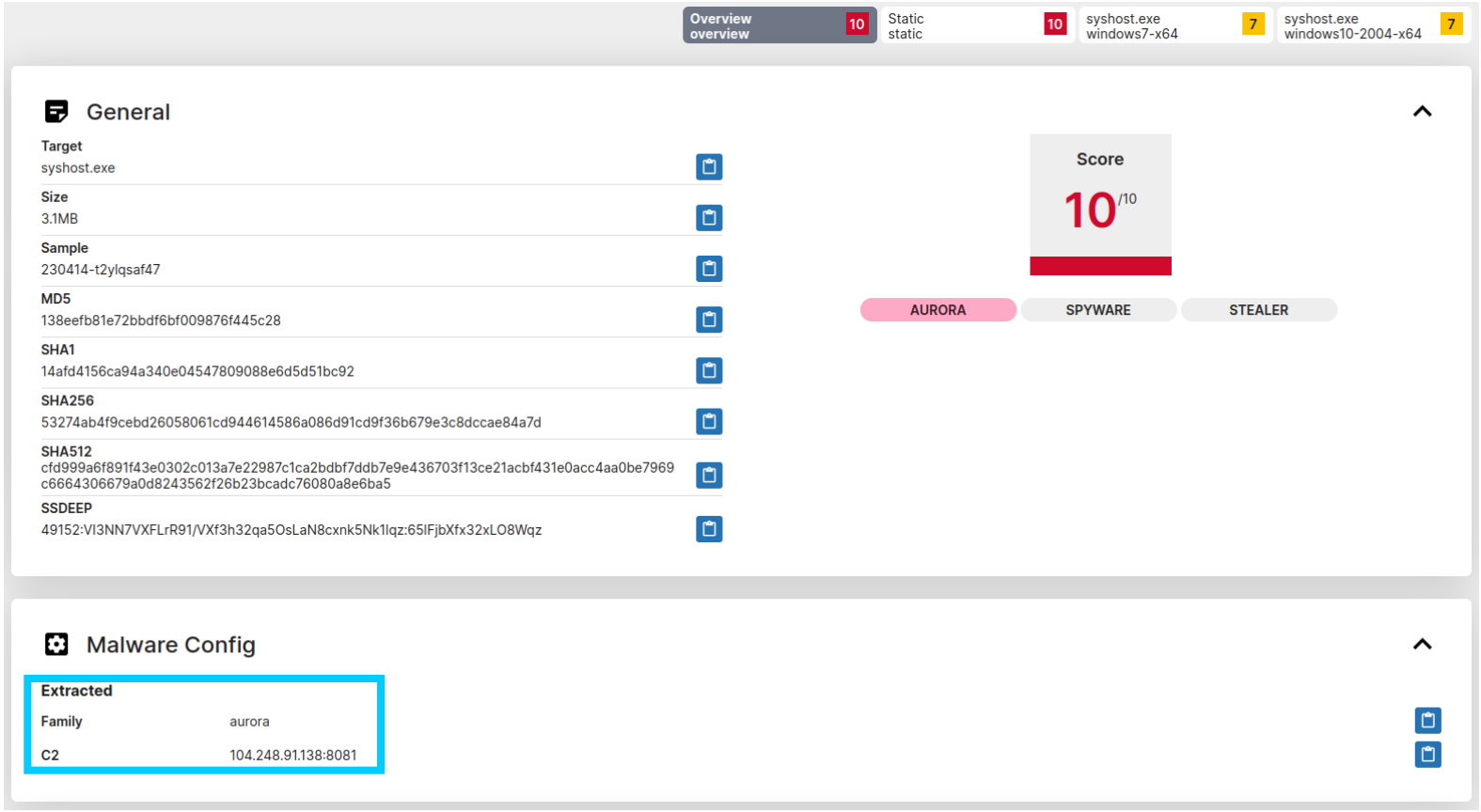
Evasion Techniques
Fallback Channel
Proxy Servers
Stepping Stones
Fast-Flux
DNS Calculation
Encryption
Data-Obfuscation
Non-Standard Ports
Traffic Pattern Manipulation

Now what?

Application of the Taxonomy

Application of the Taxonomy

Example: Aurora Stealer

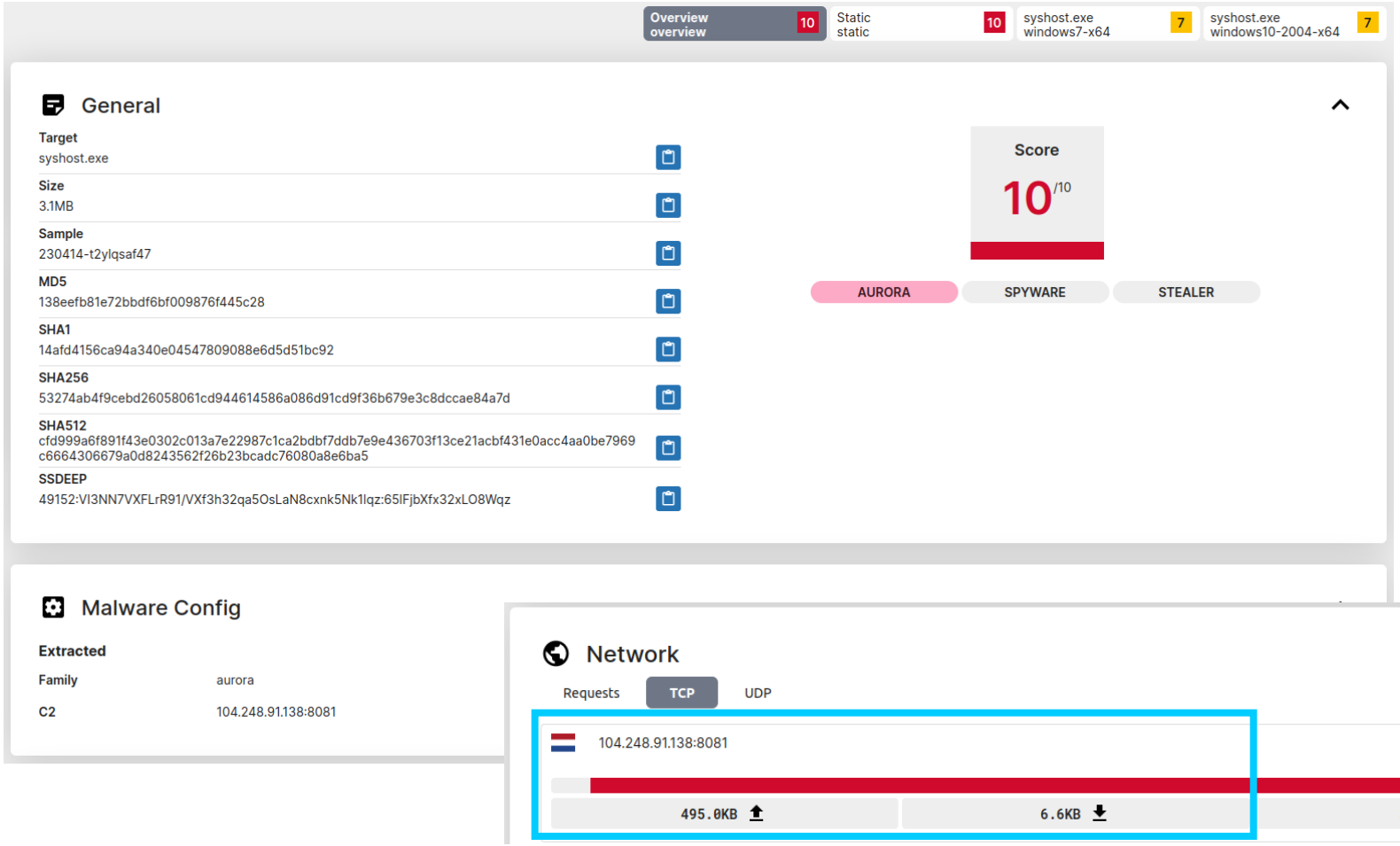


[1] https://tria.ge/230414-t2ylqsaf47/behavioral2

Fraunhofer FKIE

Application of the Taxonomy

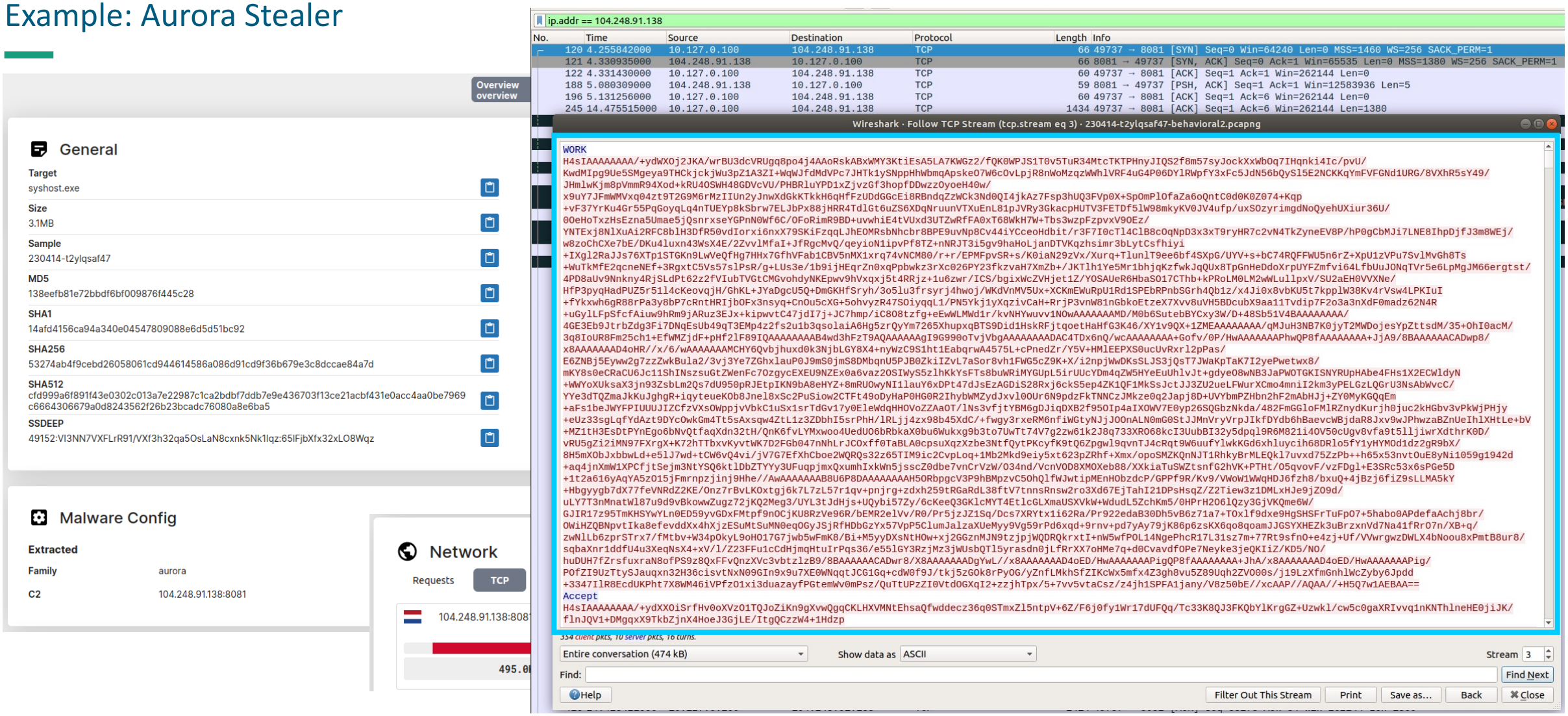
Example: Aurora Stealer



[1] https://tria.ge/230414-t2ylqsaf47/behavioral2

Application of the Taxonomy

Example: Aurora Stealer

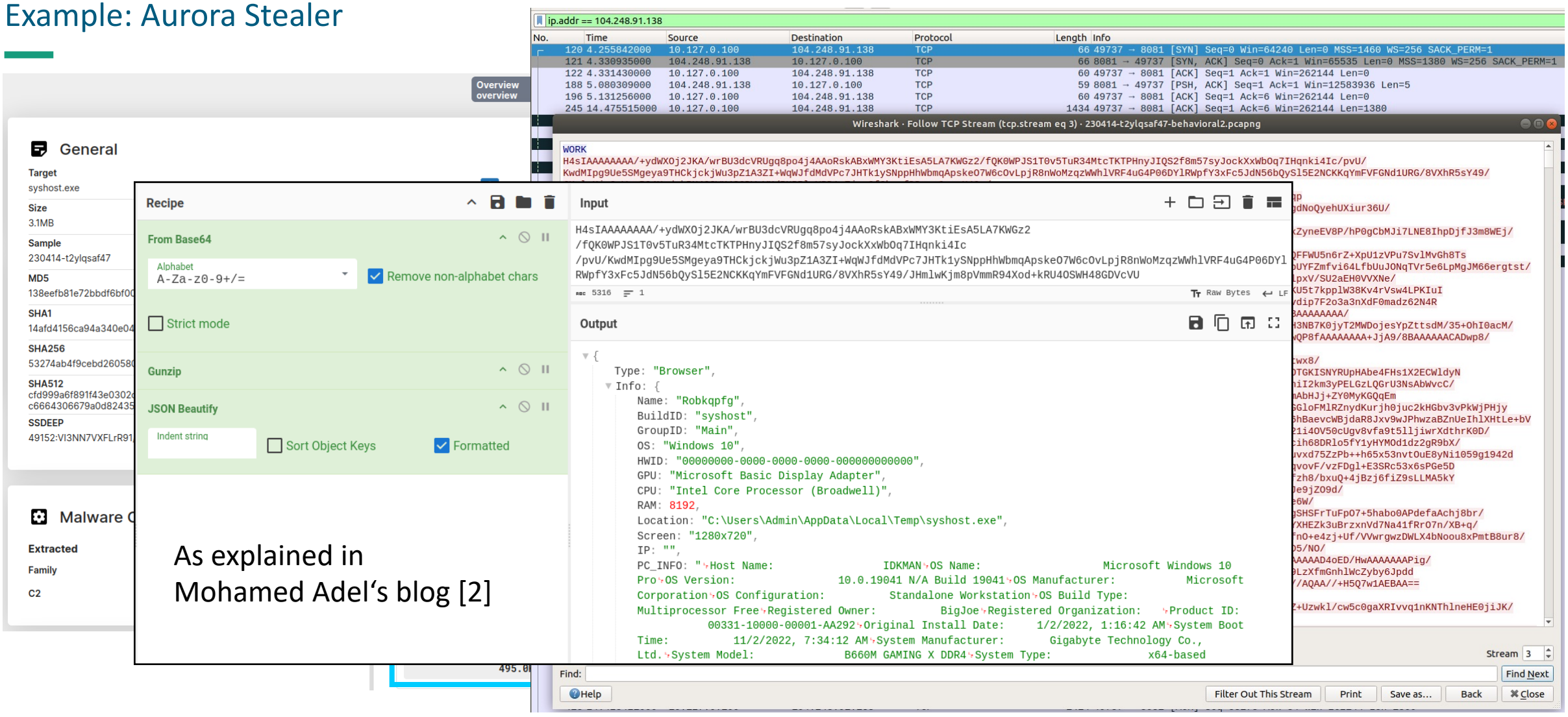


[1] https://tria.ge/230414-t2ylqsaf47/behavioral2

[2] https://d01a.github.io/aurora-stealer/

Application of the Taxonomy

Example: Aurora Stealer



[1] https://tria.ge/230414-t2ylqsaf47/behavioral2
[2] https://d01a.github.io/aurora-stealer/

Application of the Taxonomy

Example: Aurora Stealer

Application of Taxonomy Dimensions:

C&C Model	**centralized**
C&C Rallying	**hardcoded IP (1)**
Communication Behavior	**2-way PUSH**
Carrier Protocol	**raw TCP**
(Internal) C&C Protocol	**server: plain, bot: base64(gzip(json))**
Evasion Techniques	**(data obfuscation / non-standard port)**

[1] https://tria.ge/230414-t2ylqsaf47/behavioral2
[2] https://d01a.github.io/aurora-stealer/

Application of the Taxonomy

Goals

- **Verify applicability of the taxonomy**
- **Overview of malware C2 communication strategies used in current malware**

Application of the Taxonomy
Approach

- **Identify prevalent malware families**
  - **Focus on most common submissions to Malware Bazaar**
  - **Period: Oct 2022 – Feb 2024**
- **Collect reference samples and sandbox reports with live C2 traffic**
  - **Examine Tria.ge sandbox runs**
  - **Filter PCAPs and extract C2 characteristics**
- **Data Set to be published with the paper**

Application of the Taxonomy

Identification of Prevalent Malware families (Oct 2022 – Feb 2024)

- **193,920 submissions (167,755 identified by 809 different signatures)**

Rank	Family	Subm	Perc	cumPerc
1	RedLineStealer	21758	12.97	12.97
2	Mirai	20175	12.03	25.00
3	AgentTesla	20028	11.94	36.94
4	Amadey	9386	5.60	42.53
5	Smoke Loader	8915	5.31	47.84
6	Formbook	8396	5.00	52.85
7	GCleaner	6716	4.00	56.85
8	SnakeKeylogger	5015	2.99	59.84
9	RemcosRAT	3329	1.98	61.83
10	Loki	3070	1.83	63.66

Rank	Family	Subm	Perc	cumPerc
11	GuLoader	2957	1.76	65.42
12	Socks5Systemz	2895	1.73	67.15
13	Gafgyt	2822	1.68	68.83
14	CoinMiner	2168	1.29	70.12
15	Adware.Neoreklami	1810	1.08	71.20
16	njrat	1780	1.06	72.26
17	Heodo	1752	1.04	73.30
18	DCRat	1750	1.04	74.35
19	AsyncRAT	1693	1.01	75.36
20	Quakbot	1686	1.01	76.36

Application of the Taxonomy

Collect reference samples and sandbox reports with live C2 traffic

elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys
elf.mirai	win.azorult	win.darkcomet	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws	win.pikabot	win.smokeloader
win.agent_tesla	win.brute_ratel_c4	win.dbatloader	win.lumma	win.privateloader	win.socks5_systemz
win.amadey	win.bumblebee	win.dcrat	win.masslogger	win.qakbot	win.stealc
win.asyncrat	win.cloudeye	win.emotet	win.nanocore	win.quasar_rat	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire	win.recordbreaker	win.vidar
Reviewed 2200+ sandbox runs for these families in order to find runs with live traffic **Reported characteristics are as „observed“**		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader
		win.glupteba	win.panda_stealer	win.remcos	win.xworm

Application of the Taxonomy

Collect reference samples and sandbox reports with live C2 traffic

elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys
elf.mirai	win.azorult	win.darkcomet	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws	win.pikabot	win.smokeloader
win.agent_tesla	win.brute_ratel_c4	win.dbatloader	win.lumma	win.privateloader	win.socks5_systemz
win.amadey	win.bumblebee	win.dcrat	win.masslogger	win.qakbot	win.stealc
win.asyncrat	win.cloudeye	win.emotet	win.nanocore	win.quasar_rat	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire	win.recordbreaker	win.vidar
„Impure“ for most loaders (incl. loads) (5) Only runs with unresponsive servers for win.formbook and win.lokipws :(		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader
		win.glupteba	win.panda_stealer	win.remcos	win.xworm

Application of the Taxonomy

Taxonomy Dimension: C&C Model

elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys
elf.mirai	win.azorult	win.darkcomet	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws	win.pikabot	win.smokeloader
win.agent_tesla	win.brute_ratel_c4	win.dbatloader	win.lumma	win.privateloader	win.socks5_systemz
win.amadey	win.bumblebee	win.dcrat	win.masslogger	win.qakbot	win.stealc
win.asyncrat	win.cloudeye	win.emotet	win.nanocore	win.quasar_rat	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire	win.recordbreaker	win.vidar
All considered families use a centralized model **Hierarchical structure are opaque to observation** **P2P reportedly used for IoT botnets [1]**		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader
		win.glupteba	win.panda_stealer	win.remcos	win.xworm

[1] https://blog.netlab.360.com/p2p-botnets-review-status-continuous-monitoring/

Application of the Taxonomy

Taxonomy Dimension: C&C Rallying – as observed

elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys
elf.mirai	win.azorult	win.darkcomet (DDNS)	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws (IP + Domain)	win.pikabot	win.smokeloader
win.agent_tesla	win.brute_ratel_c4	win.dbatloader (OneDrive)	win.lumma	win.privateloader	win.socks5_systemz
win.amadey	win.bumblebee	win.dcrat (pastebin)	win.masslogger	win.qakbot	win.stealc
win.asyncrat	win.cloudeye (GoogleDocs)	win.emotet	win.nanocore (DDNS)	win.quasar_rat	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire (DDNS)	win.recordbreaker	win.vidar (telegram)
Hardcoded C2 info widely used **IP (23½) / Domain (19½) / SMTP (4) / Indirection (2)** **BumbleBee turned towards DGA in 2023**		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader
		win.glupteba	win.panda_stealer	win.remcos (DDNS)	win.xworm

[1] https://blog.netlab.360.com/p2p-botnets-review-status-continuous-monitoring/

Application of the Taxonomy

Taxonomy Dimension: Communication Behavior – as observed

elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys*
elf.mirai	win.azorult	win.darkcomet	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws	win.pikabot*	win.smokeloader
win.agent_tesla	win.brute_ratel_c4*	win.dbatloader	win.lumma	win.privateloader*	win.socks5_systemz
win.amadey	win.bumblebee*	win.dcrat	win.masslogger	win.qakbot*	win.stealc
win.asyncrat	win.cloudeye	win.emotet	win.nanocore	win.quasar_rat*	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire	win.recordbreaker	win.vidar
Most C2 protocols are 2-way pull (41) or push (3) **LOTS-based Loaders are considered 1-way pull (2) SMTP-based exfil is 1-way push (4)**		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader*
		win.glupteba	win.panda_stealer	win.remcos*	win.xworm*

Application of the Taxonomy

Taxonomy Dimension: Carrier Protocol – as observed

elf.bashlite 8722	win.ave_maria 5200	win.cryptbot 80	win.hawkeye_keylogger 587	win.phorpiex 80	win.rhadamanthys 443
elf.mirai 6666	win.azorult 80 CF	win.darkcomet 1604	win.isfb 80 CF	win.photoloader 80	win.sectop_rat 15647
win.404keylogger 587	win.bazarbackdoor 9001	win.darkgate 80	win.lokipws 80	win.pikabot 2967	win.smokeloader 80
win.agent_tesla 587	win.brute_ratel_c4 443	win.dbatloader 443	win.lumma 443 CF	win.privateloader 27323	win.socks5_systemz 80
win.amadey 80	win.bumblebee 443	win.dcrat 80	win.masslogger 587	win.qakbot 443,2222, 995, …	win.stealc 80
win.asyncrat 9999	win.cloudeye 443	win.emotet 80	win.nanocore 5899	win.quasar_rat 4412	win.tofsee 443
win.aurora_stealer 8081	win.cobalt_strike 443	win.formbook 80	win.netwire 3102	win.recordbreaker 80	win.vidar 9100
Raw TCP (11) (TLS (3)) HTTP (18) (TLS (14)) SMTP (4)		win.gcleaner 80	win.njrat 13538	win.redline_stealer 4132	win.wikiloader 443
		win.glupteba 443	win.panda_stealer 80	win.remcos 2718	win.xworm 65030

[1] https://blog.netlab.360.com/p2p-botnets-review-status-continuous-monitoring/

Application of the Taxonomy

Taxonomy Dimension: (Internal) C&C Protocol – as observed

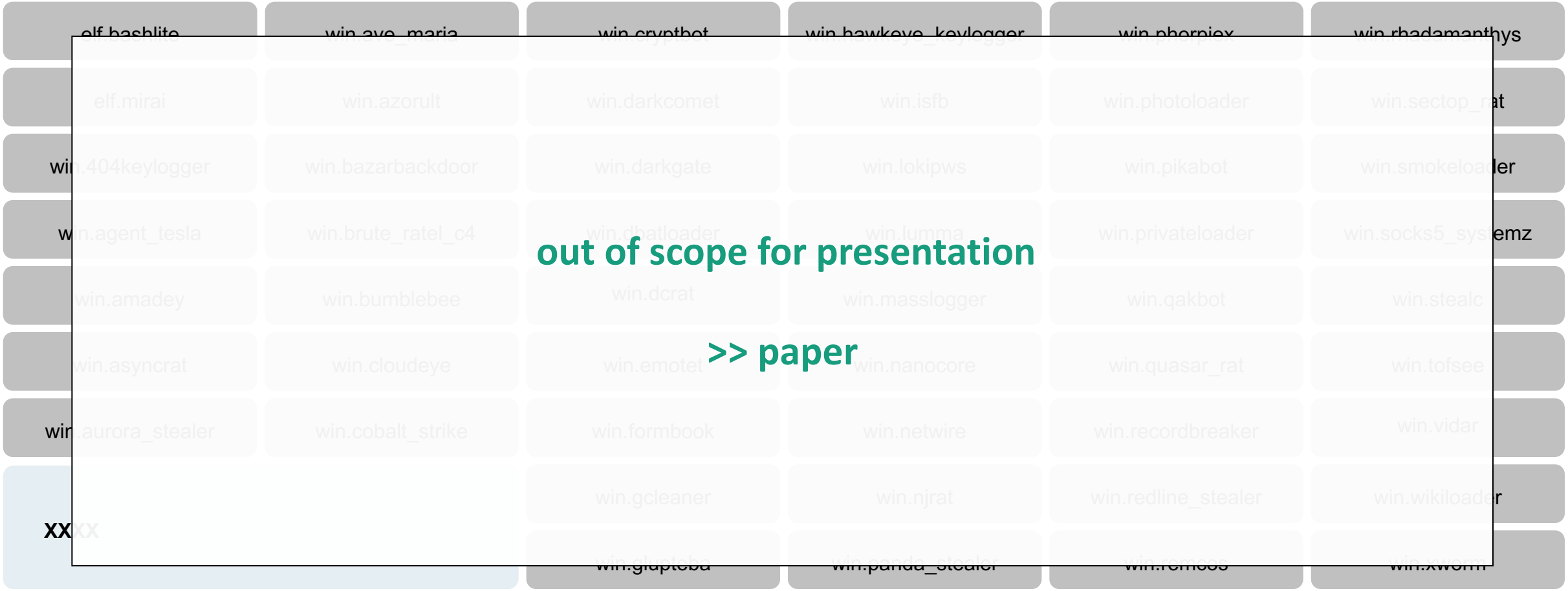
elf.bashlite	win.ave_maria	win.cryptbot	win.hawkeye_keylogger	win.phorpiex	win.rhadamanthys
elf.mirai	win.azorult	win.darkcomet	win.isfb	win.photoloader	win.sectop_rat
win.404keylogger	win.bazarbackdoor	win.darkgate	win.lokipws	win.pikabot	win.smokeloader
win.agent_tesla	win.brute_ratel_c4	win.dbatloader	win.lumma	win.privateloader	win.socks5_systemz
win.amadey	win.bumblebee	win.dcrat	win.masslogger	win.qakbot	win.stealc
win.asyncrat	win.cloudeye	win.emotet	win.nanocore	win.quasar_rat	win.tofsee
win.aurora_stealer	win.cobalt_strike	win.formbook	win.netwire	win.recordbreaker	win.vidar
		win.gcleaner	win.njrat	win.redline_stealer	win.wikiloader
		win.glupteba	win.panda_stealer	win.remcos	win.xworm

Text-based (9) (encrypted (1))
Binary (5) (encrypted (1))
Unknown – encrypted (17)
Carrier-protected (17)

[1] https://blog.netlab.360.com/p2p-botnets-review-status-continuous-monitoring/

Application of the Taxonomy

Taxonomy Dimension: Evasion Techniques



out of scope for presentation

>> paper

[1] https://blog.netlab.360.com/p2p-botnets-review-status-continuous-monitoring/

Fraunhofer FKIE

Conclusion

Conclusion

1. Taxonomy

C&C Models	
Centralized	Star
Centralized	Hierarchical
P2P	
Hybrid	
Random	

Rallying Mechanisms	
IP-Address	Hardcoded
IP-Address	Seeding
Domain Names	Hardcoded
Domain Names	DGA
Further Indirection	

Communication Behavior	
Transmission	Push/Pull
Sessions	1-way/2-way

C&C Protocols
Custom Text-Based
Custom Binary
Existing Protocols
...

Carrier Protocols
Raw TCP/UDP Socket
Raw SSL/TSL Socket
Application Layer Procols
Web Services

Evasion Techniques
Fallback Channel
Proxy Servers
Stepping Stones
Fast-Flux
DNS Calculation
Encryption
Data-Obfuscation
Non-Standard Ports
Traffic Pattern Manipulation

Now what?

2. Case-Study

elf.bashlite 8722	win.ave_maria 5200	win.cryptbot 80	win.hawkeye_keylogger 587	win.phorpiex 80	win.rhadamanthys 443
elf.mirai 6666	win.azorult 80 CF	win.darkcomet 1604	win.isfb 80 CF	win.photoloader 80	win.sectop_rat 15647
win.404keylogger 587	win.bazarbackdoor 9001	win.darkgate 80	win.lokipws 80	win.pikabot 2967	win.smokeloader 80
win.agent_tesla 587	win.brute_ratel_c4 443	win.dbatloader 443	win.lumma 443 CF	win.privateloader 27323	win.socks5_systemz 80
win.amadey 80	win.bumblebee 443	win.dcrat 80	win.masslogger 587	win.qakbot 443,2222, 995, …	win.stealc 80
win.asyncrat 9999	win.cloudeye 443	win.emotet 80	win.nanocore 5899	win.quasar_rat 4412	win.tofsee 443
win.aurora_stealer 8081	win.cobalt_strike 443	win.formbook 80	win.netwire 3102	win.recordbreaker 80	win.vidar 9100
Raw TCP (11) (TLS (3)) HTTP (18) (TLS (14)) SMTP (4)		win.gcleaner 80	win.njrat 13538	win.redline_stealer 4132	win.wikiloader 443
		win.glupteba 443	win.panda_stealer 80	win.remcos 2718	win.xworm 65030

3. Dataset