# Attacker Behavior Modeling using Temporal Data Analysis

Azqa Nadeem PhD candidate @ Cyber Analytics Lab Department of Intelligent Systems Delft University of Technology





07 June 2022

### **Dynamic observables**

- Program execution  $\rightarrow$  observable data
  - Software logs
  - Network traffic
  - Intrusion alerts
  - ...



- Shows exactly what happened (not what could've happened)
- (Relatively) obfuscation free



Proxy to attacker intent TUDelft

#### **Temporal traces**

• Temporal events  $\rightarrow$  insightful patterns

Challenges:

- Curse of dimensionality
- Visualization?
- Distance measure?
- Performance

- ...



### Attacker behavior modeling

- Attacker strategy extraction
  - Data: Intrusion alerts
  - Method: Probabilistic Deterministic Finite Automata
- Malware behavioural profiling
  - Data: Network traffic
  - Method: Hierarchical density-based clustering



## **USE CASE I: ATTACKER STRATEGY EXTRACTION**



#### Alert management is expensive

#### • Security analysts handle >1M intrusion alerts/day\* → Alert fatigue

",\"event type\":\"alert\",\"sic ip\":\"10.0.254.30\",\"src port\":59481,\"dest ip\":\"10.0.0.23\",\"dest port\":3006,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"qid\":1,\"signature id\":2010937,\"rev\":2,\"signature\":,\"EF FOLICY Suspicious inbound to mySQL port ",\"event type\":\"alert\",\"sic ip\":\"10.0.254.30\",\"src port\":59482,\"dest ip\":\"10.0.0.153\",\"dest port\":110,\"proto\":\"TCP\",\"alert\":(\"action\":\"110ved\",\"qid\":1,\"signature id\":2002992,\"rev\":7,\"signature ",\"event type\":\"alert\",\"sic ip\":\"10.0.254.30\",\"src port\":59482,\"dest ip\":\"10.0.0.190\",\"dest port\":993,\"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature id\":2002995,\"rev\":10,\"signature\":\"ET SCAN Rapid IMAPS Connections - Possibl ",\"event type\":\"alert\",\"sic ip\":\"10.0.254.30\",\"signature id\":2002511,\"rev\":6,\"signature\":"ET SCAN Potential VNC Scan 5900-5920\",\"c \"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":59481,\"dest\_ip\":\"10.0.0100\",\"dest\_port\":3889,"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"qid\":1,\"signature\_id\":2014884,\"rev\":8,\"signature\":"ET DOS Microsoft Remote Desitop (RDP) Sym ',\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":59481,\"dest\_ip\":\"10.0.0.167\",\"dest\_port\":3389,\"proto\":\"allowed\",\"gid\":1,\"signature\_id\":2001972,\"rev\":20,\"signature\":\"ET SCAN Behavioral Unusually fast Termin \"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":59482,\"dest\_ip\":\"10.0.0100\",\"dest\_port\":3889,\"proto\":\"ICP\",\"alert\":\"allowed\",\"gid\":1,\"signature\_id\":2014884,\"rev\":8,\"signature\":\"ET DOS Microsoft Remote Desitop (RDP) Sym ",\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":59483,\"dest\_ip\":\"10.0.0.110\",\"dest\_port\":995,\"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2002993,\"rev\":7,\"signature\":,"ET SCAN Rapid POP3S Connections - Possible ',\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":59481,\"dest\_ip\":\"10.0.223\",\"dest\_port\":3389,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2014384,\"rev\":8,\"signature\":\"ET DOS Microsoft Remote Desktop (RDP) Syr "\"event type\":\"alert\".\"src ip\":\"10.0.254.30\".\"src port\":59481.\"dest ip\":\"10.0.0.101\".\"dest port\":3389.\"proto\":\"TCP\".\"alert\":\"alert\":\"alert\":\"alert\":\"alert\":\"alert\":\"signature id\":2014384.\"rev\":8.\"signature\":8.\"signature\":\"ET DOS Microsoft Remote Desitop (RDP) Syn ./\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":59483,\"dest\_ip\":\"10.0.47\",\"dest\_port\":143,\"proto\":\"TCP\",\"alert\":(\"action\":\"alloved\",\"gid\":1,\"signature\_id\":2002994,\"rev\":7,\"signature\":\"ET SCAN Rapid IMAP Connections - Possible B: "\"event type\":\"alert\".\"src ip\":\"10.0.254.30\".\"src port\":59484.\"dest ip\":\"10.0.202\".\"dest port\":445 traff \"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_jo\":\"10.0.2021,"\"6est\_ip\":\"10.0.2021,","dest\_port\":135,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2001581,\"rev\":15,\"signature]":\"ET SCAN Behavioral Unusual Port 125 traffi '\\"event type\":\"alert\".\"src ip\":\"10.0.254.30\".\"src port\":59482.\"dest ip\":\"10.0.0.151\".\"ignature\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I30.0.151\".\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I30.0.151\".\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I39.\"proto\":\"I30.0.151\".\"proto\":\"I39. .\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.31\",\"src\_port\":45860,\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":43860,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2010927,\"rev\":2,\"signature\":\"ET FOLICY Suspicious inbound to mySQL port ',\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.31\",\"src\_port\":45861,\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":5900,\"proto\":\"allowed\",\"gid\":1,\"signature\_id\":2002911,\"rev\":6,\"signature\":\"ET SCMN Potential VNC Scan 5900-5920\",\" ("event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45861,\"dest ip\":\"10.0.99.16\",\"dest port\":47EP\";(\"action\":\"TCP\",\"alert\";(\"action\":\"TCP\",\"alert\";(\"action\":\"TCP\",\"alert\";(\"action\":\"TCP\",\"alert\";(\"action\":\"TCP\";)"] .\"event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45860,\"dest ip\":\"10.0.99.245\",\"dest port\":3306,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature id\":2,\"signature .\"event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45860,\"dest ip\":\"10.0.99.245\",\"dest port\":22,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"ginature id\":2001219,\"rev\":20,\"signature\":\"ET SCAN Fotential SSH Scan\",\"category\": ',\"event\_type\":\"alert\",("src\_ip\":\"10.0.254.30\",\"src\_port\":65046,\"dest\_ip\":\"10.0.222\",\"dest\_port\":2389,\"proto\":\"TCP\",\"alert\";(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2014384,\"rev\":8,\"signature\":\"ET DOS Microsoft Remote Desitop (RDP) Syr \\"event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45862,\"dest ip\":\"10.0.99.16\",\"dest port\":45862,\"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature id\":2002910,\"rev\":6,\"signature\":"ET SCAN Potential VNC Scan 5800-5820\",\"c .\"event type\":\"alert\",\"src ip\":\"10.0.254.30\",\"src port\":65046,\"dest ip\":\"10.0.0.71\",\"dest port\":5900,\"proto\":\"TCP\",\"alert\":("action\":\"allowed\",\"gid\":1,\"signature id\":2002911,\"rev\":6,\"signature\":\"ET SCAN Potential VNC Scan 5900-5920\",\"c: ',\"event type\":\"alert\",\"src ip\":\"10.0.254.30\",\"event type\":\"10.0.223\",\"dest ip\":\"10.0.223\",\"dest port\":142,\"proto\":\"TCP\",\"alert\":("action\":\"allowed\",\"gid\":1, "signature id\":200294,\"rev\":7, "signature\":\"ET SCAN Rapid IMAP Connections - Possible ",\"event type\":\"alert\",\"sic ip\":\"10.0.254.30\",\"sic port\":65046,\"dest ip\":\"10.0.0.23\",\"dest port\":2306,\"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"qid\":1,\"signature id\":2010937,\"rev\":2,\"signature\":,\"EF FOLICY Suspicious inbound to mySQL port .\"event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45860,\"dest ip\":\"10.0.99.44\",\"dest port\":1433,"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"qid\":1,\"signature id\":2010925,\"rev\":2,\"signature\":\"EF FOLICY Suspicious inbound to MSSQL port ",\"event type\":\"alert\",\"signature id\":2002995,\"rev\":65046,\"dest ip\":\"10.0.77\",\"dest port\":992,\"proto\":\"TCP\",\"alert\":("alert\":("alert\",\"ainature id\":2002995,\"rev\":10,\"signature\":\"ET SCAN Rapid IMAPS Connections - Possible \"event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45860,\"dest ip\":\"10.0.99.44\",\"dest port\":458(0,\"dest p ',\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.30\",\"src\_port\":65059,\"eev\":10.0.0.81\",\"dest\_port\":45,\"proto\":\"TCP\",\"alert\":("alert\",\"alert\":("alert\",("aignature\_id\":2001569,\"rev\":15,\"signature\":\"ET SCAN Behavioral Unusual Port 445 traffi ",\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.31\",\"src port\":45860,\"dest\_ip\":\"10.0.99.44\",\"dest\_port\":1521,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2010936,\"rev\":2,\"signature\":2,\"signature\":\"ET FOLICY Suspicious inbound to Oracle S( .\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.35\",\"src\_port\":33314,\"dest\_ip\":\"10.0.99.44\",\"dest\_port\":3306,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2010937,\"rev\":2,\"signature\":\"ET FOLICY Suspicious inbound to mySQL port "\"event type\":\"alert\".\"src ip\":\"10.0.254.35\".\"src port\":33316.\"dest ip\":\"10.0.99.16\".\"groto\":\"TCP\".\"alert\":\"alert\" ',\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.35\",\"src\_port\":33225,\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":22,\"proto\":\"TCP\",\"alert\":("alored\",\"gid\":1,\"signature\_id\":2001219,\"rev\":20,\"signature\":\"ET SCAN Potential SSH Scan\",\"category\": ',\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.35\",\"src\_port\":23315,\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":2306,\"proto\":\"TCP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2010927,\"rev\":2,\"signature\":\"ET POLICY Suspicious inbound to mySQL por .\"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.31\",\"src\_port\":45860,\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":423,\"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature\_id\":2010925,\"rev\":2,\"signature\":\"ET FOLICY Suspicious inbound to MSSQL port .\"event type\":\"alert\",\"src ip\":\"10.0.254.31\",\"src port\":45860,\"dest ip\":\"10.0.99.16\",\"dest port\":5432,\"proto\":\"TCP\",\"alert\":{\"action\":\"allowed\",\"gid\":1.\"signature id\":2010939,\"rev\":2.\"signature\":\"ET POLICY Suspicious inbound to PostgreSOI "\"event type\":\"alert\".\"src ip\":\"10.0.254.31\".\"src port\":45860.\"dest ip\":\"10.0.99.16\".\"dest port\":1521,\"proto\":\"TCP\".\"alert\":(\"action\":\"allowed\".\"did\":1.\"signature id\":2.\"signature\":2.\"signature\":2.\"signature\":2.\"signature\":\"ET POLICY Suspicious inbound to Oracle SC \"event\_type\":\"alert\",\"src\_ip\":\"10.0.254.35\",\"src\_port\":32314,\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":5802,\"proto\":\"TCP\",\"alert\":\"alert\":\"alert\",\"src\_ip\":\"10.0.954.35\",\"src\_ip\":\"10.0.99.16\",\"dest\_ip\":\"10.0.99.16\",\"dest\_port\":5800-5820\",\"c \"event type\":\"alert\",\"src ip\":\"10.0.254.35\",\"src port\":33314,\"dest ip\":\"10.0.99.44\",\"dest port\":1433,\"proto\":\"ICP\",\"alert\":(\"action\":\"allowed\",\"gid\":1,\"signature id\":2010935,\"rev\":2,\"signature\":\"EF POLICY Suspicious inbound to MSSOL port

**ŤU**Delft



https://itsec.group/blog-post-csoc-guide.html

### Attack graphs

• "Attack Graphs" display attacker strategies



### Alert-driven Attack graphs

- How to discover and display attacker strategies?
- SAGE: IntruSion alert-driven Attack Graph Extractor





### Anatomy of an Alert-driven Attack Graph







#### 1. Alert-type imbalance





- 1. Alert-type imbalance
- 2. Context matters

**∦ RIT** TUDelft



Scan<sub>1</sub>, Scan<sub>2</sub>, ACE, Exfiltration, ... vs. Scan<sub>1</sub>, Scan<sub>1</sub>, Scan<sub>2</sub>, Scan<sub>2</sub>, ...



- 1. Alert-type imbalance
- 2. Context matters
- 3. Explainable approach







### SAGE: IntruSion alert-driven Attack Graph Extractor





#### Alerts → Episode/Action sequences

{	'_sourcetype 'alert': {	': 'suricata:alert', 'category': 'Attempted Information Leak', 'severity': 2.									
		'signature': 'ET POLICY Python-urllib\\/ 'Suspicious User Agent'},									
	'dest_ip': '169.254.169.254', 'dest_port': 80, 'src_ip': '10.0.0.20', 'src_port': 56952.										
	'timestamp': '2018-11-03T13:51:58.205548+0000'}}										

 $C \leftrightarrow D$ 



Hosts:10.0.254.206->10.0.0.20











### Suffix-based PDFA

- Suffix-based Probabilistic
  Deterministic Finite Automaton
- Summarizes attack paths
- Brings infrequent episodes to the top
   Red → Severe | Blue → Medium severity





### Suffix-based PDFA

- State identifiers model context
- Markovian properties
  - States  $\rightarrow$  milestones with context
- Low-severity Sinks ignored
  - Cleaner model





### Adding context & Attack graph formation



#### **Experimental dataset**

- Suricata alerts from Collegiate Penetration Testing Competition<sup>1</sup>
  - 6 multi-attacker teams
  - 1 fictitious network
  - 330,270 alerts

**ŤU**Delft

- Moskal's Action-Intent framework<sup>2</sup>
  - Alert signature  $\rightarrow$  Attack stage



CPTC dataset: <u>https://www.globalcptc.org/</u>

2. S. Moskal and S. J. Yang, "Framework to describe intentions of a cyber attack action," arXiv preprint arXiv:2002.07838, 2020.

### [1] Alert triaging

- 330,270 alerts  $\rightarrow$  93 alert-driven AGs
- Compresses ~500 alerts in < 25 vertices

	# alerts (raw)	# alerts (filtered)	#episodes	#ES/ #ESQ	#ESS	#AGs
<b>T1</b>	81373	26651	655	103	108	53
<b>T2</b>	42474	4922	609	86	92	7
<b>T5</b>	52550	11918	622	69	74	51
<b>T7</b>	47101	8517	576	63	73	23
<b>T8</b>	55170	9037	439	67	79	33
<b>T9</b>	51602	10081	1042	69	110	30



### [2] Attacker strategy visualization

- Shows how the attack transpired
- 3 teams, 5 attempts

**ŤU**Delft

- 3 ways to reach objective
  - Discovered by S-PDFA





### SAGE: Summary & Future work

- SAGE uses sequence learning for automated strategy extraction
  - S-PDFA critical for accentuating infrequent events & modeling context
- Alert-driven attack graphs
  - Compress thousands of alerts in a few AGs
  - Provide insight into attacker strategies and behavior dynamics

- Attack path prioritization
- Missing paths in AGs
- Modeling evolving strategies **TUDelft** Alert-driven attack graphs using S-PDFA. Azqa Nadeem, Sicco Verwer, Stephen Moskal, Shanchieh Jay Yang. In IEEE Transactions on Dependable and Secure Computing, 2021.

## USE CASE II: MALWARE BEHAVIOR PROFILES



#### Inconsistent malware family labels

• Malware labels are inconsistent and black-box





#### AV vendors & their naming conventions

Dridex-Loader -	-	-	-	-	-	-	-	-	-	-	-	3	12
ZeuS-OpenSSL -	-	-	-	-	-	-	-	-	-	-	-	1	1
DridexRAT -	-	-	-	-	-	-	-	-	-	-	-	-	7
Dridex -	-	-	-	-	-	-	-	-	-	-	-	3	3
ZeuS-VM-AES -	-	-	-	-	-	-	-	10	-	-	-	15	4
ZeuS -	-	-	-	-	-	-	-	-	-	-	-	3	-
Gozi-ISFB -	-	-	-	14	6	-	-	-	45	-	-	37	20
Zeus-Panda -	-	-	-	-	-	-	-	-	-	-	-	10	-
Ramnit -	-	-	-	-	-	-	12	-	-	-	-	3	7
ZeuS-v1 -	-	-	-	-	-	-	-	-	-	-	-	10	-
Zeus-Action -	-	-	-	-	-	-	-	-	-	-	-	2	-
Blackmoon -	77	-	700	-	-	31	-	41	-	-	11	11	16
Gozi-EQ -	-	-	-	-	-	-	-	-	-	-	-	7	-
ZeuS-P2P -	-	-	-	-	-	-	-	-	-	-	-	4	-
Citadel -	-	1	-	-	-	-	-	-	-	26	-	12	31
	banbra -	citadel -	dinwod -	gamarue -	- gozi	qzonit -	ramnit -	- razy	ursnif -	zbot -	- Ksnz	OTHERS -	SINGLETON -

Vendor 2

Vendor 1



### Malware behavior profiles

- Malware labels are inconsistent and black-box
- Behavior profiles are more insightful of capabilities





#### Malware capability assessment

- How to discover behaviors and build profiles?
- MalPaCA: Malware Packet Sequence Clustering and Analysis



Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families. Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.

#### Network trace collection

FOX IT

• Malware infected machine generates network traffic

<u> </u>	No.	Source	Destination	Protoc	Lengtł	Info
		40 192.168.1.2	192.168.1.110	ICMP	82	Redirect (Redirect for host)
		41 CzNicZSP_00:0	PcsCompu_7c:9	ARP	60	192.168.1.1 is at d8:58:d7:00:0f:72
		42 192.168.1.110	203.153.165.21	TCP	182	49191 → 8343 [PSH, ACK] Seq=1 Ack=1 Win=65700 Len=128
		43 203.153.165.21	192.168.1.110	TCP	60	8343 → 49191 [ACK] Seq=1 Ack=129 Win=15744 Len=0
		44 203.153.165.21	192.168.1.110	TCP	1188	8343 → 49191 [PSH, ACK] Seq=1 Ack=129 Win=15744 Len=1134
		45 192.168.1.110	203.153.165.21	ТСР	380	49191 → 8343 [PSH, ACK] Seq=129 Ack=1135 Win=64564 Len=326
		46 192.168.1.2	192.168.1.110	ICMP	408	Redirect (Redirect for host)
		47 203.153.165.21	192.168.1.110	TCP	113	8343 → 49191 [PSH, ACK] Seq=1135 Ack=455 Win=16768 Len=59
		48 fd2d:ab8c:225	fd2d:ab8c:225	DNS	110	Standard query 0xb554 A www.download.windowsupdate.com
		10 100 100 1 110	000 4F0 40F 04	TOD	F 4	

**FUDE Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families.** Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.

#### Network trace collection

• Malware infected machine generates network traffic



lo.	Source	Destination	Protoc	Lengtł	Info
40	192.168.1.2	192.168.1.110	ICMP	82	Redirect (Redirect for host)
41	CzNicZSP_00:0	PcsCompu_7c:9	ARP	60	192.168.1.1 is at d8:58:d7:00:0f:72
42	192.168.1.110	203.153.165.21	TCP	182	49191 → 8343 [PSH, ACK] Seq=1 Ack=1 Win=65700 Len=128
43	203.153.165.21	192.168.1.110	TCP	60	8343 → 49191 [ACK] Seq=1 Ack=129 Win=15744 Len=0
44	203.153.165.21	192.168.1.110	TCP	1188	8343 → 49191 [PSH, ACK] Seq=1 Ack=129 Win=15744 Len=1134
45	192.168.1.110	203.153.165.21	ТСР	380	49191 → 8343 [PSH, ACK] Seq=129 Ack=1135 Win=64564 Len=326
46	192.168.1.2	192.168.1.110	ICMP	408	Redirect (Redirect for host)
47	203.153.165.21	192.168.1.110	TCP	113	8343 → 49191 [PSH, ACK] Seq=1135 Ack=455 Win=16768 Len=59
48	fd2d:ab8c:225	fd2d:ab8c:225	DNS	110	Standard query 0xb554 A www.download.windowsupdate.com

Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families. Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.

#### **Behavior catalog construction**

**FOX IT** 



**FUDE Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families.** Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.

### Similarity analysis

• Distance calculation with distortions in sequences



#### **Dynamic Time Warping**



### Behavior (Cluster) analysis



**FUDEIft** Temporal Heatmaps for Visualizing Sequential Features. <u>https://cyber-analytics.nl/blogposts/2020-12-28-temporal-heatmaps/</u>

#### **Malware Behavior Profiles**

		В	С	D	DL	GE	GI	R	Z	ZP	ZPa	Zv1	ZVA
	SSDP traffic	1	✓	1	1	1	1	1	1	-	✓	-	1
	Broadcast traffic	1	1	-	1	-	~	1	-	~	-	1	1
	LLMNR traffic	1	1	-	1	-	~	-	-	-	-	-	-
	System. port scan	1	1	-	-	-	~	1	-	-	-	-	1
	Random. port scan	1	1	-	-	-	~	~	-	-	-	-	1
Benavior	In conn spam	-	-	-	-	-	~	-	-	-	-	-	-
cataloa	Out conn spam	-	-	-	-	-	~	-	-	-	-	-	-
	Malicious Subnet	-	-	-	-	-	-	-	-	-	-	-	1
	In HTTPs	-	1	-	1	-	~	-	-	-	1	-	-
	Out HTTPs	-	-	-	-	-	~	-	-	-	1	-	-
	C&C reuse	1	-	-	-	-	-	-	-	-	1	-	-
	Misc.	1	✓	-	1	-	1	-	1	-	1	-	~
FOX I	# Clusters	7	11	1	8	1	16	4	2	1	7	1	7

**FUDELT** Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families. Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.

#### **Malware Behavior Profiles**



Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families. Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.

### MalPaCA: Summary & Future work

- Malware family names  $\rightarrow$  noisy & inconsistent
- MalPaCA for building behavioral profiles
  - Clustering multivariate packet sequences
  - Clusters → Malware behavior catalog
  - Malware behavioral profile  $\rightarrow$  cluster membership

- Network + system behavior catalog
- Continual learning
- Adversarial robustness

Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families. Azqa Nadeem, Christian Hammerschmidt, Carlos H. Ganan, Sicco Verwer. In Malware Analysis using Artificial Intelligence and Deep Learning, Springer, 2021.



- Goal: Learning attacker behavior from temporal data
- Unsupervised setting with limited prior knowledge
- Input: Observables | Output: Intelligence

- 2 use-cases
  - Intrusion alerts  $\rightarrow$  Attacker strategies via attack graphs
  - Network traffic  $\rightarrow$  Malware behavior profiles



## **Questions**?

Goal: Learning attacker behavior from temporal data Unsupervised setting with limited prior knowledge Input: Observables | Output: Intelligence

2 use-cases Intrusion alerts → Attacker strategies via attack graphs Network traffic → Malware behavior profiles

azqa.nadeem@tudelft.nl

https://cyber-analytics.nl/

Icons courtesy Freepik, Those Icons, Dinosoft Labs, Pixel Perfect, Flat Icons, Kiranshastry from www.flaticon.com