A semantic approach for describing Advanced Persistant Threat

A. Berady - G. Guette - M. Jaume - V. Viet Triem Tong

EPI CIDRE INRIA / CentraleSupelec / CNRS / Univ. Rennes 1 valerie.viettriemtong@centralesupelec.fr

> Supsec Winter Workshop January 2023





APT : Advanced Persistant Threat

A term coined by Colonel Greg Rattray (US Air Force) in 2006 and popularized the NIST in 2011 $\,$

The Advanced Persistent Threat :

- pursues its objectives repeatedly over an extended period of time;
- adapts to defenders' efforts to resist it; and
- is determined to maintain the level of interaction needed to execute its objectives.

Before 2011, the real knowledge of APTs remains confidential When the term APT started to be used, the general public has heard about

- Moonlight Maze (1996) : targeting US military and government networks pointing to Russian Internet Service Providers in 1996
- Titan Rain (2003) series of attacks in the US since 2003 originated from China
- StuxNext (2010) uncovered in 2010 and thought to have been in development since at least 2005, widely understood to be a

cyberweapon against Iran

Supsec Winter Workshop A semantic for APT description

 Operation Aurora (2010) series of cyber attacks originated from China targeting over 20 US companies





More than 10 years later, if you want to study APT?

Few datasets [1]

- I won't talk about KDD99
- Unified Host and Network Dataset [2]
- DAPT 2020 [3]
- PWNJUTSU 2022 [4]

Some un-structured reports

AptNotes https://github.com/aptnotes/ Operation Aurora, Malware Targeting Organizations in Ukraine

Videos, tweet and other media

• TV5 Monde

Few (No ?) details on the targeted architecture, the defense system, the precise attack scenario

First Step : Global overview Lifecycle of an Advanced Persistant Threat

APT Lifecycle : Cyber Kill Chain Lockheed Martin in 2011

- Linear model focusing on the initial compromise
- Cannot describe long-term attacks



ATP-Life-cycle : Kill chains models

Pols in 2017 Unified Kill Chain

- introduces the notion of repetitiveness of technical actions
- introduces the notion of phases of APT
- does not consider the potential regression of the attacker.



Tactics of the Attack matrix

MITRE ATT&CK in 2013 a knowledge base of TTPs.

MITRE ATT&C	K'					Matrice	a Tactica * Te	chriques * Data Sources	Mitgations * 0	roups Software	Compaigns Resour	ces + Blog (?	Contribute Search Q,
Reconnaissance	Resource	Initial Access	Execution	Persistence	Privilege Escelation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Command and Control	Extination	Impact
10 techniques	7 techniques	S techniques	13 sechniques	19 techniques	13 techniques	42 techniques	17 techniques	30 techniques	5 techniques	17 techniques	14 techniques	5 techniques	13 techniques
Active Scanning _{CE}	Acquire infrastructure _{Ch}	Drive-by Compromise	Command and Scripting Intersection up	Account Internet on	Abuse Elevation Centrol Mechanismus	Abuse Devertion Control Mechanism v.	Adversarpin-the- Mode or	Account Discovery (4	Exploitation of Remote Zervices	Adversery in the Moder or	Application Layer	Automated Editorial co	Account Access Remeval
Cother Victim Heat Information (K	Compromise Accounts (p)	Exploit Public Facing Application	Container Administration Command	BITS ARE	Access Tokes Manipulation on	Access Taken Manipulation _{(C}	Brute Farce (0)	Application Window Discovery Browser Boolemark Discovery	Indexed Spearphishing	Author Collected Detw (1)	Communication Through Removable Media	Data Transfer Blor Um/to	Data Destruction Data Encrypted for Impact
Eather Victors Identity Information of	Compromise Infrastructure (1)	Edenai Renate Services	Deples Centainer	Direct or Legen Autoritat: Direction (14)	Scot or Logen Autoster1	872.364	Credentials from Percentral Stores on	Coul Infrastructure Discovery	Laberal Tool Transfer	Aude Ceptere	Data Encoding ₂₀	Enflowion Over	Data Maripulation or
Cether Victors Network	Develop Capabilities ₂₀	Kantware Additions	Exploitation for Client	Rect or Legre	Execution (14)	Build shage an road	Exploritation for	Cost Service Dashboard	Session Hipdoing 11	Automated Collection	Data Obluscation (1)	Protocol (8	Delecement pp
And and a second second	Establish Accounts 20	Prohing on	Execution	Constantion Screen ap	Initialization Scripts ₂₁	Debugger Exaster	Creating Access	Dout Service Discovery	Remote Services (4	Accessor Senators	Dynamic Resolution ₁₀	Extension Over E2	Disk Work 10
Information (4	Obtain Capabilities _N	Repleator Trough	Communication (8	Concernancial Chart	Create or Modily	Information	Forces Percenting and	Cloud Storage Object Discovery	Replication Trengh	- Haland Care	Trengint Channel ()	Common Com Select	Empairs Denial of
Phibling for Information 30	Shape Capabilities (6)	Survey Chain	Native AP1	Self-nere Brury	homain Enders	Deploy Container	Credentials (8	Container and Resource	Software Engloyment	Data from Claud	Fallback Channels	Setwork Medium (1)	Deneral Constrine
Search Dosed Services (8)		Comprender (4	Scheduled Tesh/Jok (4	Create Account pp	Medification (c)	Direct Volume Access	Input Capture (K)	Patroner Lension	Taula	Donage	Ingress Taol Transfer	Editorian Over Druging Medican	Inhibit Sustem Excessory
Beandy Open Technical Detailognet up		Trusted Belationship	Seveless Execution	Create or Madily System Process or	Escape to Heat	Corrain Policy Medification (2)	Modly Adherituation Process	Jonah Trat Bacavery	Tabl Shared Conterd	Data from Configuration	Multi-Stage Charvels	Editoria Over Nob	Network Central of
Search Open		Fabil Accounts (4	Shared Modules	Event Trippered	Event Toppered Evention we	Execution Daardrafs (r)	Matifactor	File and Directory Discovery	Use Alternate Appendication	Repository (3	Nan-Application Laper Protocol	Service (4	Service pt
Reduited/Domains pa			Baltware Deployment Taolo	Constant (red	Delotation for Privilege	Exploitation for Defense Evenior	Authendication Interception	Group Policy Discovery	Maria (k	Bata from Information Repositories (2)	Non-Standard Port	Exheduled Standler	Resource Hijsching
Nobshee			System Services (2	Services	Excession	File and Directory Permissions	Multifactor .	Network Service Discovery		Data from Local System	Prefocol Tanneling	Cloud Account	Deves The
			User Decution (1)	Hist Decution	Pervoya	Modification pa	Seneration Require	Natwork Share Discovery		Sata from Network	Press (4		Spillers Shuldown/Micout
			Windows Management	- Prof 10	Process Injection (25	Part America	Network Snilling	Natwork Cellfing		And the form the stability	Remote Access		
				Interfact the second strapp	Scheduled Text/Job (5	house belocase	05 Credential	Pesseneral Paring Discovery		Media	Table Section		
				Process (1	Yalid Accounts (A	Indicative Descent	Stand Analyzation	Periphenel Device Discovery		Sate Steped 15	and Service -	1	
				Office Application Startup as		Indirect Command Execution	Access Tokes	Permission Groups Discovery on		Email Collection (6			
				Pre-05 Bost.m.		Macquerading -s	Deal or Purge Authentication	Process Discovery		Input Capture (4)	•		
				Scheduled Task/Job ps		Modily Authentication	Cetificates	Cuery Repirity		Screen Capture			
				Server Software		Pracess py	Takets _M	Remote System Discovery		Video Capitare			
				Camponens pp		inhearty close compose	Steel Neb Seesion	Software Discovery (5					
				Contra angenerry py		Modify Reporty	- Cooks	System Information Discovery					
						Modify System Image 13	Credentials (5	System Lecetion Discovery (c)					
						Network Boundary Bridging (1)		System-Network Configuration					
						Obtacated Files or Information (p)		Suran Network Connections					
						Plat File Modification		System Duries User Discovery					
						Pre-OS-Boot ₍₃₎		System Service Discovery					
						Process Injection (FD)		Bystem Time Discovery					
						Rogue Borrein Centroller		Whatston Tankos Design of					

Tactics of the Attack matrix

MITRE ATT&CK is not a model per se but it deepens the notion of phase of an attack without highlighting their ordering



Modeling the Operational Phases of APT Campaigns [5]



Modeling the Operational Phases of APT Campaigns [5]



Modeling the Operational Phases of APT Campaigns [5]



Instantiation by the incident of TV5 Monde





Suppose that the community agrees on a generic model to represent an APT

but we still lack data ..

varied, representative, up-to-date and above all accurate data



PWNJUTSU project

?WNJUTSU

- Project funded and supported by IRSN BCyP
- 22 professional attackers attacks on a dedicated architecture
- New available dataset !

Publication

Aimad Berady, Mathieu Jaume, Valérie Viet Triem Tong et Gilles Guette : PWNJUTSU: A Dataset and a Semantics-Driven Approach to Retrace Attack Campaigns.

IEEE Transactions on Network and Service Management (TNSM), Special Issue on Recent Advances in Network Security Management, 2022.

PWNJUTSU Project overview

- 3 machines (Windows and Linux) : $M_1
 ightarrow M_2
 ightarrow M_3$
- mandatory checkpoints with *flags* to recover
- Several attack paths
- Vulnerabilities easy to exploit, so that the experimentation is focused on propagation in the network.





PWNJUTSU Project – Overview

- Dedicated Instances for each participant.
- Probes on operating systems and verbose logs
- Continuous capture of network flows
- Supervision by a SIEM.



Figure – PWNJUTSU infrastructure

PWNJUTSU Project – Participants



- 22 experts from the TOP100 of YesWeHack experts
- 9 nationalities
- Progressive and attractive financial rewards
- Typical participant profile :
 - 25-35 years old (63%);
 - Bachelor's/Master's level degree (91%);
 - Certified in "ethical hacking". (64%);
 - Self-trained offensive security expert(100%).

PWNJUTSU dataset

https://pwnjutsu.irisa.fr

- a raw dataset
 - 16 million system events
 - 172 GB of network traffic
 - a search engine

21/	H-A		netstat 1			
			All players	- Syden aduly	- Query substitution	
i UL	DL					
			3. 29-09/TPORT	E S. HAM SOM	3. BURETEN	0. 10401
SERVER Facepage August.	The online news	dan kela kena	us the first 52 vesualis. Field free to deves	shoud it is get all the data.		
Time 1	Playar	Level	Trace			
3/05/2821 11:02:48	009	and	May 12 11/02/48 n/19-ym1 anos	py[2528]: (aid:0-aid:2498-try/dev(pts/0-cwd/h	arae/han_acia filenarae/bis/ <mark>retatet</mark>): retatat
5/05/2821 08:48:53	n21	sen1	May 15 00:42:53 n21 (vm1 anos)	py(3054): (ald:1120 ald:3080 tty:/dev/pts/1 cw	t/home/lando, calification filename./l	in/netstat) netstat -
5/05/2821 08:45:09	n21	and .	May 15 00.45:09 n21 ym1 anos	py(3122): (aid:1120 aid:3097 ttp:/dev/pts/1 ov	t/home/lando,calitizaian filename.4	on/netstat) netstat -
6/05/2821 17:14:50	n23	sen1	May 10 17:14:30 n23 rml snort	py[1637]: [aid:1124 sid:857 ity (nene) ovd./tp/	/readme.opp filenome./bin/ nebslat.	entriet Hump
6/05/28211818-59	n23	sen1	May 10 10:10:09 n23 yml snos	py(1812): (aid:1127 sid:1798 Hy./dev/pts/1 ew	dultome/Zsec Nenemeubin/ <mark>netatat</mark>	drut- tateter
6/05/28211811.44	123	with 1	May 16 18 11 44 n23 west second	ov[1828] fails at 1798 systematic out or	ot filenerve.tbir/ netstet) netstet -	turip
5/05/2821 28:09:05	125	west 1	May 19 20.09.08 r28 with second	ogliticij facti sati 1870 ily decistari ovezv	ot forware.ton/ netalet) netalet -	tanep
5/05/2821 28 21 49	125	sen1	May 19 23 21 49 r25 with soon	07.001 (01.001 1120 641 20309 10, /0ev/pix/0	wd/http://encourt.com/MMXX) are	NR - NO
5/05/2821 28:21.49	126	and	May 19 23 21:49 r26-yen1 anon	04(16703) (via 1120 via 26339 m/view/pro/0	wit/http:fiesame/bit/metatal) are	NIM - Mpl
1.05/2821 12.39.48	127	and	May 21 12/39/43 rg7 wint store	ogli 101 janu sati 2654 ny/secipti 20 cont/1	lename/bir/retstat) retstat -ani	

Here we have data

but how to present them?

we need a way to detail the whole scenario and each particular attack progression

Attacker's report



Informel report where some element are missing, attacker's perspective only

P12 progression



This progression has been manually inferred and represented.

An attacker centric model to retrace attack campaign



An attacker state

- μ an attack position (machine, user)
- $\bullet \ \mathcal{S}$ the recovered secrets
- E a partial view of the targeted system

A targeted system is a set of machines. A machine **m**

- \mathbb{S}_m : services
- $\bullet \ \mathbb{P}_m$: some files
- A_m : accounts
- $\mathbb{N}_{\mathbf{m}}$: a neighboring

Progression of an attacker

A complete attack campaign is a sequence of attacker states representing the evolution of his control of the target.

The attacker moves from one state to another by applying an attack technique.

$$(\mu_i, \mathcal{S}_i, \mathcal{E}_i) \xrightarrow{\mathbf{t}(params)} (\mu_{i+1}, \mathcal{S}_{i+1}, \mathcal{E}_{i+1})$$

The attack techniques are those defined by the MITRE attack These techniques are still defined in nature language and do not have a precise semantic.

MITRE ATT&CK - T1210

MITRE | ATT&CK

Exploitation of Remote Services

Adversaries may exploit remote services to gain unauthorized access to internal systems once inside of a network. Exploitation of a software vulnerability occurs when an adversary takes adventage of a programming error in a program, service, or within the operating system software or kernel itself to execute adversary-controlled code. A common goal for post-compromise exploitation of remose services is for latent invorvement to exelute access to a remote system.

An adversary may need to determine if the remote system is in a vulnerable state, which may be done through Network. Service Discovery or other Discovery methods looking for common, vulnerable software that may be deployed in the network, the lack of eratin patches that may indicate submeabilities, or security software that may be used to detect or contain remote exploitation. Servers are likely a high value target for lateral movement exploitation, but endpoint systems may also be at risk if they privoke an dwarting or access to additional resources.

There are several well-known vulnerabilities that exist in common services such as SMB ^[1] and RDP ^[2] as well as applications that may be used within internal networks such as MySQL ^[3] and web server services.^[4]

Depending on the permissions level of the vulnerable remote service an adversary may achieve Exploitation for Privilege Escalation as a result of lateral movement exploitation as well.

ID: T1210

Sub-techniques: No sub-techniques

- Tactic: Lateral Movement
- i Platforms: Linux, Windows, macOS
- ③ System Requirements: Unpatched software or otherwise vulnerable target. Depending on the target and goal, the system and exploitable service may need to be remotely accessible from the internal network.
- Permissions Required: User

Contributors: ExtraHop

Version: 1.1

Created: 18 April 2018

Last Modified: 24 February 2022

Version Permalink

Procedure Examples

ID	Name	Description
G0007	APT28	APT28 exploited a Windows SMB Remote Code Execution Vulnerability to conduct lateral movement.[5]0[7]
S0606	Bad Rabbit	Bad Rabbit used the EternalRomance SMB exploit to spread through victim networks. ^[8]
S0608	Conficker	Conficker exploited the MS08-067 Windows vulnerability for remote code execution through a crafted RPC request. ^[9]

A semantic for the technique Exploitation of Remote Services

Technic	T ₁₂₁₀ : Exploitation of Remote Services
TACTIC	Lateral movement
Description	Gain access to a machine by remotely exploiting a vulnerability
	using x exploit on an exposed network service s.
PARAMETERS	m, u, m', s, x
Préconditions	$(m,u)\in\mu$,
	$\mathbf{m}' \in \lfloor \mathbb{N}_{\mathbf{m}} \rfloor_{\mathcal{E}},$
	$s \in \lfloor \mathbb{S}_{\mathbf{m}'} \rfloor_{\mathcal{E}}$ et
	$x \in Exploits(s)$
TRANSITION	$(\mu, \mathcal{S}, \mathcal{E}) \hookrightarrow (\mu', \mathcal{S}, \mathcal{E})$
	where $\mu' = \mu \cup \{(\mathbf{m}', \mathbf{u}')\}$
	with $(u', s, k, \ell) \in \mathbb{A}_{m'}$
VARIANTS	Authenticated vulnerabilities use the additional parameters $\mathbf{u}^{\prime\prime}$ and
	k'' such as $(\mathbf{u}'', s, k'', \ell'') \in \lfloor \mathbb{A}_{\mathbf{m}'} \rfloor_{\mathcal{E}}$

A semantic for the technique Network Service Scanning T1046

Technique	T ₁₀₄₆ : Network Service Scanning
TACTIQUE	Discovery
Description	Discover all network services of a remote machine \mathbf{m}' by browsing
	the namespace of network ports $\Delta \subseteq \{0, \cdots, 65535\}$.
Paramètres	m, u, m', Δ
Préconditions	$(m,u)\in\mu$,
	$m' \in \lfloor \mathbb{N}_{m} \rfloor_{\mathcal{E}}$
TRANSITION	$(\mu,\mathcal{S},\mathcal{E}) \hookrightarrow (\mu,\mathcal{S},\mathcal{E}')$ with
	$ \mathcal{E}' = \mathcal{E} \left[\mathbf{m}' \leftarrow \left(\lfloor \mathbb{S}_{\mathbf{m}'} \rfloor_{\mathcal{E}} \cup \{ \mathbf{s}(\text{port}:i) \mid i \in \Delta \}, \lfloor \mathbb{P}_{\mathbf{m}'} \rfloor_{\mathcal{E}}, \lfloor \mathbb{A}_{\mathbf{m}'} \rfloor_{\mathcal{E}}, \lfloor \mathbb{N}_{\mathbf{m}'} \rfloor \right) $

Attack technics semantics

In [4] we detail the specification of 13 techniques, which satisfy 5 tactics :

- Lateral Movement : horizontal movement in the network (same user, different machine);
- Credential Access : collection of credentials;
- **Privilege Escalation** : vertical movement in the network (different user, same machine);
- Discovery : discovery of the technical environment;
- **Persistence** : implementation of a permanent remote access mechanism.

P12 progression (reminder)

Extract from P12 report

- scan nmap (1000 ports) through the VPN.
- 2 Discovery of several services.
- 3 Recover banners and discover the continuum application.
- Launched a bruteforce on the SSH port (without success and not very functional).
- Search for public vulnerabilities on continuum.
- Usage of Metasploit module to successfully exploit the continuum vulnerability.
 - I obtained a shell and fast environment of the machine.



Evolution of Player 12's knowledge



The complete attack campaign of Player 12

Step 0	P12 got an initial access to n12-gateway.
See 1	P12 performed activork service scanning (T1046) from
owy :	n12-gateway to n12-vml.
Parameters	M = n12 - gateway, U = anonymous
Truce (not)	W = h12 - VH1, A = {coproceportannap}
	40 42548 + 445 (598) Sepri Mis-1024 Les-0 #56-1357
Step 2	P12 exploited remote service (T1210) Apache Continuum (port
Presentation	SUBLO from n12-gateway to n12-yml.
	M' = n12 - vml
	S = continuum(port: 8080)
Trace (net)	K = EUB-10 : 3 9 94 5" 170911 2021-09-09 20:28:15.498303 172.24.128.112 10.12.1.1
	#TTP 1411 POST /montinuum/savwisataliation.action HTTP/1.1
Step 3	P12 got a secret flag file (T1083) on n12 - vn1.
FUNDERD	<pre>M = hi2 - vh1, 0 = root D =/opt/apache continuum//flag.txt</pre>
Trace (sys)	May 9 28128154 nl2-vml ansopy[1566]:
	filename:/bis/cati: cat flag.cot
Step 4	P12 got all credentials (T1003.008) of n12 - vn1 OS.
Parameters	M = m12 - vml, U = root, S = OS _{lives}
110.0 (SJX)	<pre>Map w servers nurves manpy(1999): [widt1 widt1009 ttp:/scom/ codi/opt/apachs_continues/</pre>
	filenamer/bis/catl: cat /#tc/sbadow
auto a	P12 added a private key (from the root user's folder) for the user has sola on the service SSH (T1136) on machine n12 = vm1.
Parameters	M=m12-vml, u=root,
	U'= han_solo, K = SSHprivatekey,
Trace (sys)	Rev 9 2013616 millioni manerol18331
	(width midt1000 ttp:/sone) cwdt/root filesamer/min/mv/t
Sec. 6	P12 discovered remote system (T1018) o12-sym2 from
	n12-vm1 using ARP table.
Parameters	M = n12 - vml, u = root, M = n12 - vml
ince (iyi)	Nay 9 31(39(3) 012-vel mmopy[1423): [aid:0 aid:1309 Uty:[score] cuds/home/has_scio/.ssh filenames/ans/mbin/asy]:
	arp -an
A DESCRIPTION OF A DESC	· · · · · · · · · · · · · · · · · · ·
Step 7	P12 performed network service scanning (T1046) from n12-vm1 to n12-vm2 as user root.
Step 7 Parameters	P12 performed network service scanning (T1046) from n12-vml to n12-vm2 as aser root. M = n12 - vm1, U = root.
Step 7 Parameters	P12 performed network service scanning (T1046 from n12-vml to n12-vml, das user root. M = n12 - vml, u = root M' = n12 - vml, u = root M' = n12 - vml, u = root
Step 7 Parameters Trace (sys)	$\label{eq:products} \begin{array}{l} P(2) gentomed served serves scanning (T1046) from n12-vrl the n12-vrl as user root.\\ \hline M = n12 - vrl u = root.\\ \hline M = n12 - vrl u = root.\\ \hline M = n12 - vrl u = root.\\ \hline m_{2} = 12 (vrl u = 1, 0) (0) (0) (0) (vrl u = 1, 0) (0) (0) (0) (0) (0) (0) (0) (0) (0) $
Step 7 Parameters Trace (sys)	$\label{eq:produced servers scaning (T1046 from n12-vml lb n12-vml s stor root. \dot{M} = n12 - vml, u = root. \\ \dot{M} = n12 - vml, u = n12 $
Step 7 Parameters Trace (sys) Step 8	$ \begin{array}{l} \texttt{Pl}2 \mbox{ problems} d \mbox{ envises sensing}(T1046 from $n12$-$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
Step 7 Parameters Trace (sys) Step 8 Parameters	$ \begin{array}{l} \text{PI2} potomod served kerne kanning (TIM46 from n12-m21 to 12-m21 to$
Step 7 Parameters Trace (sys) Step 8 Parameters	PT2:protocols serves sussing (TDIAM) from n12-vn1 in n12-vn1 are root. M 1217 - vn1, un = root. M 2217 - vn1, un = root
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys)	$\label{eq:response} \begin{split} & proves service (TOMAG free nl2-vn2 is nr2-vn2 is response to resolve the resolve service (TOMAG free nl2-vn2 is response to resolve service) (TOMAG free nl2-vn2 is response to resolve service) (TOMAG free nl2-vn2 is response to resolve service) (TOMAG free nl2-vn2 is resolved and resolve service) (TOMAG free nl2-vn2 is resolved and resolved service) (TOMAG free nl2-vn2 is resolve$
Step 7 Parameters Trace (194) Step 8 Parameters Trace (195) Year 8	If 2 photometa served, serves, saming (TDMS from n.1.2-vs.) We satisfy a server of the serves of th
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9	The photometer derives across series (TDMA from n.2.—or) The approximation of the series (TDMA from n.2.—or) M = n.2.[m], as the point of the series (TDMA from n.2.—or) M = n.2.[m], as the point of the series (TDMA from n.2.—or) TDMA from the series of the protocols about the series (TDMA from the brown one from a noise that (SDMA (TDMA from the brown of the series (TDMA from the brown
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9	$\label{eq:constraints} \begin{array}{l} \text{Prior} I = \text{prior}(B) & \text$
Step 7 Parameters Trace (sys) Step 8 Parameters Step 9 Parameters	The photons denois strength constraints of the first strength of the strength
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9 Parameters	The photon schemes are set of 1000 from (-1) ==1. We have a scheme strength of the photon scheme strength of the photon scheme strength of the photon scheme sch
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9 Parameters Trace (aut)	The definition between the second se
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9 Parameters Trace (set) Step 10	The definition denotes the second se
Step 7 Parameters Trace (syl) Step 8 Parameters Trace (syl) Step 9 Parameters Trace (set) Step 10	$\label{eq:second} \begin{array}{l} \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\$
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9 Parameters Trace (ast) Step 10 Step 11	The properties between the second se
Step 1 Parameters Trace (sys) Step 8 Parameters Trace (set) Step 10 Step 10	The definition between the second se
Step 7 Parameters Tince (1931) Step 8 Parameters Tince (1931) Step 9 Parameters Tince (1931) Step 10 Step 11 Step 12	$ \begin{array}{l} \begin{array}{l} \\ \label{eq:second} \\ \label{eq:second} \\ \\ \label{eq:second} \\ \\ eq$
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9 Parameters Trace (set) Step 10 Step 11 Step 12 Step 12	 The definition derives are seen as the second second
Step 7 Parameters Trace (sys) Step 8 Parameters Trace (sys) Step 9 Parameters Trace (set) Step 10 Step 11 Step 12 Step 12 Step 14	The definition between the second se
Step 7 Parameters Tince (sys) Step 8 Parameters Tince (sys) Step 9 Parameters Tince (set) Step 10 Step 11 Step 14	The second seco
Step 1 Parameters Take (sys) Step 8 Parameters Tike (sys) Step 9 Parameters Tiker (sys) Step 9 Parameters Tiker (sys) Step 10 Step 11 Step 13 Step 14	The operation of the second se
Step 7 Parameters Tince (sys) Step 8 Parameters Tince (sys) Step 9 Parameters Tince (stel) Step 10 Step 11 Step 13 Step 13	The definition denotes the second se
Step 7 Parameters Thace (oys) Step 8 Parameters Tince (oys) Step 9 Parameters Tince (oss) Step 10 Step 11 Step 11 Step 13 Step 13 Step 15 Step 15	The second secon
Step 1 Parameters Time (sys) Step 8 Parameters Time (sys) Step 9 Parameters Time (sys) Step 9 Parameters Time (set) Step 10 Step 11 Step 12 Step 13 Step 14 Step 15 Step 16	The definition denotes one constrained of the definition of the d
Step 7 Parameters Tince (sys) Step 8 Parameters Tince (sys) Step 9 Parameters Tince (ste) Step 10 Step 11 Step 13 Step 13 Step 14 Step 13 Step 14 Step 17	$\label{eq:2.1} \begin{tabular}{lllllllllllllllllllllllllllllllllll$

- 18 steps
- 6 attack techniques used
- 6 attack positions

First immediate benefit

You can visualize the attack from the attack techniques point of view $% \left({{{\mathbf{r}}_{\mathrm{s}}}_{\mathrm{s}}} \right)$



First immediate benefits

You can visualize the propagation area



Perspectives

Attack scenario

- An attack position is a pair (machine, user)
- A successful attack procedure execution
 - increase the attacker knowledge
 - or allows to move from an attack position to another



Take away

- PWNJUTSU : a new dataset of traces of professional attackers
- a semantic of attack techniques that allows to precisely describe the attacker behavior
- the central concept of attack position

What is still missing?

In this work

- the dataset misses from noise and normal activity
- the logs were manually interpreted

More globally

We need

- more precisely described datasets with different infrastructures and different attacks
- high level and low level representation of attacks
- a way to infer these representation automatically

Perspectives

Submitted for publication. PhD thesis research project of Pierre-Victor Besson

Generation of training systems



Perspectives

submitted for publication, to be improved during the PhD thesis research project of Manuel Poisson in collaboration with Amossys.

Evaluation of the propagation area using Living-off-the-land techniques



References I

Branka Stojanović, Katharina Hofer-Schmitz, and Ulrike Kleb. Apt datasets and attack modeling for automated detection methods : A review. *Comput. Secur.*, 92, 2020.

Catherine Beazley, Karan Gadiya, Ravi K U Rakesh, David Roden, Boda Ye, Brendan Abraham, Donald E. Brown, and Malathi Veeraraghavan. Exploratory data analysis of a unified host and network dataset. In *2019 Systems and Information Engineering Design Symposium (SIEDS)*, pages 1–5, 2019.

Sowmya Myneni, Ankur Chowdhary, Abdulhakim Sabur, Sailik Sengupta, Garima Agrawal, Dijiang Huang, and Myong Kang. Dapt 2020 - constructing a benchmark dataset for advanced persistent threats. In Deployable Machine Learning for Security Defense - 1st International Workshop, MLHat 2020, Proceedings, Communications in Computer and Information Science. Springer Science and Business Media Deutschland GmbH, 2020.

Aimad Berady, Mathieu Jaume, Valérie Viet Triem Tong, and Gilles Guette. PWNJUTSU : A dataset and a semantics-driven approach to retrace attack campaigns.

IEEE Transactions on Network and Service Management, 2022.

References II

৾৾



Aimad Berady, Valérie Viet Triem Tong, Gilles Guette, Christophe Bidan, and Guillaume Carat. Modeling the Operational Phases of APT Campaigns. In CSCI 2019 - 6th Annual Conf. on Computational Science & Computational Intelligence, Las Vegas, United States, December 2019.