Sistematizarea Analizei Dinamice a Malware-ului: Dezvoltarea unei metode de clasificare a mostrelor malware prin analiză dinamică automată

Mircea Zeiconi[1] Ciprian Pungilă[2]

[1]Master Student, Faculty of Mathematics and Computer Science, West University of Timișoara,
[2]Supervisor: Conf. univ. dr. Ciprian Pungilă, West University of Timișoara

Introducere

În contextul amenințărilor cibernetice tot mai sofisticate, analiza comportamentală a fișierelor malware a devenit esențială pentru identificarea și prevenirea atacurilor informatice. Tehnicile clasice de detecție, bazate pe semnături, sunt depășite în fața variantelor avansate care folosesc mecanisme de evaziune, persistență și injectare de cod.

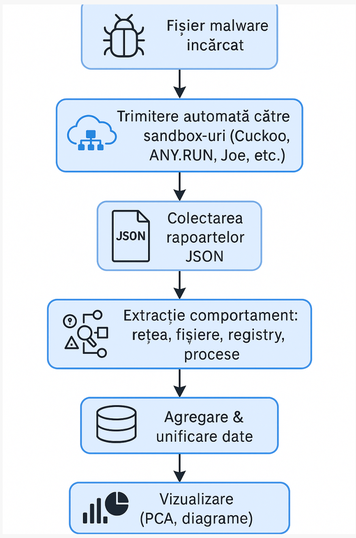
În urma analizei realizate, lucrarea propune următoarele contribuții practice:

- Automatizarea analizei comportamentale pentru fișiere malware.
- Utilizarea multiplă a sandbox-urilor (Cuckoo, ANY.RUN, Joe, etc.).
- Colectarea și unificarea rezultatelor în format JSON.
- Clasificare automată și vizuală pe baza comportamentului observat.
- Detectarea tehnicilor de evaziune, persistență și injecție.

Obiective

- Identificarea limitărilor metodelor clasice de detecție a malware-ului.
- Clasificarea malware-ului pe baza comportamentului observat în sandbox-uri.
- Automatizarea procesului de analiză pe multiple platforme.
- Colectarea și consolidarea rezultatelor într-o structură unificată.
- Evaluarea acurateței clasificării și a capacității de detecție vizuală (ex: PCA).

Flux de lucru al analizei comportamentale a fișierelor malware



Fluxul evidențiază etapele esențiale în analiza automată a fișierelor malware: încărcarea mostrei, execuția în sandbox-uri multiple, colectarea rapoartelor JSON, extragerea comportamentului, agregarea datelor și vizualizarea pentru clasificare.

Rezultatele analizei comportamentale în sandbox-uri

Sandbox	Scor	Fișiere observate	Indicatori de rețea
Cuckoo	10	2 fișiere	1 domeniu, trafic HTTP
ANY.RUN	100	2 fișiere	HTTP, HTTPS, IP-uri externe
JoeSandbox	95	2 fișiere	DNS request, comandă C2
HybridAnalysis	100	2 fișiere	Verdict: *Malicious*
Triage	9	3 fișiere	Conexiuni C2, DNS
Detux	—	—	Incompatibil cu Windows (.exe)

Tabelul sumarizează scorul comportamental, activitatea pe fișiere și indicatorii rețelei pentru fiecare sandbox.

Vizualizare rezultate – PCA (dimensionalitate redusă)

Reducerea dimensionalității folosind PCA

- Separare clară între fișierele malware și cele benigne.
- Grupări vizibile în spațiul bidimensional generat prin PCA.
- Util pentru clasificare automată și interpretare comportamentală.

Semnături comportamentale detectate

Semnătură comportamentală	Sandbox-uri
Creare cheie de autorun în registry	Cuckoo, JoeSandbox
Copiere fișier în Startup	ANY.RUN, HybridAnalysis
Execuție comenzi via `cmd.exe`	JoeSandbox, Triage
Injecție în `svchost.exe`	JoeSandbox, HybridAnalysis
Contact cu server C2	Toate
Creare fișiere temporare	Triage, Cuckoo
Modificare registry pentru persistență	HybridAnalysis, JoeSandbox

- Semnăturile evidențiază comportamente specifice fișierelor malware.
- Detectarea lor permite clasificarea și corelarea între platforme.
- Indicii pentru persistență, evaziune și injecții de cod.

Output JSON – Rezumat comportamental

Fișier	Scor	Semnături comportamentale
`sample1.exe`	10	Creează cheie de autorun, copie în Startup, contactează server C2
`dropper.tmp`	100	Injectare în `svchost.exe`, execuție `cmd.exe`, modificare registru pentru persistență
`payload.doc`	95	Deschide proces `cmd.exe`, accesează domeniu malițios, elimină copii shadow

Concluzii

- Automatizarea procesului de analiză comportamentală crește eficiența în detectarea malware-ului necunoscut.
- Clasificarea bazată pe comportament permite diferențierea clară între fișierele benigne și cele malițioase.
- Vizualizarea cu PCA oferă suport interpretabil pentru analiza datelor și detectarea anomaliilor.

Lucrări viitoare

- Integrarea unor modele de învățare automată (Random Forest, SVM).
- Analiza comportamentală pe fișiere Office sau scripturi.
- Detectarea automatizată a familiilor malware prin clustering.

Referințe

[1] Manuel Egele, Theodoor Scholte, Engin Kirda, and Christopher Kruegel.
A survey on automated dynamic malware-analysis techniques and tools.
ACM Computing Surveys (CSUR), 44(2):1–42, 2012.

[2] Arash Habibi Lashkari and David Gil.
Bodmas: A behavior-based malware analysis dataset.
In *2020 International Carnahan Conference on Security Technology (ICCST)*, pages 1–8. IEEE, 2020.

[3] ResearchGate.
Automated malware behavior analysis using json reports, 2025.

Acces suplimentar – Cod QR

Pentru acces rapid la proiectul complet, demonstrație sau codul sursă, scanează codul QR care accesează: `https://github.com/mirceazeiconi/Cod-QR`

Codul oferă acces la:

- Pagina Github cu codul Python și datele JSON;
- Documentația completă a analizei;