UNIVERZITA PAVLA JOZEFA ŠAFÁRIKA V KOŠICIACH
PRÍRODOVEDECKÁ FAKULTA

BEHAVIORÁLNE ASPEKTY VYBRANÝCH ČASTÍ RANSOMWARE

Bakalárska práca

2018 **Júlia Kázsmérová**

UNIVERZITA PAVLA JOZEFA ŠAFÁRIKA V KOŠICIACH
PRÍRODOVEDECKÁ FAKULTA

BEHAVIORÁLNE ASPEKTY VYBRANÝCH ČASTÍ RANSOMWARE

Bakalárska práca

Študijný program:	Informatika/Aplikovaná informatika
Študijný odbor:	9.2.1. Informatika
Školiace pracovisko:	Ústav informatiky
Vedúci práce:	RnDR. JuDr. Pavol Sokol, PhD.
Konzultant:	Mgr. Ladislav Bačo

Košice 2018 **Júlia Kázsmérová**

Univerzita P. J. Šafárika v Košiciach
Prírodovedecká fakulta

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Júlia Kázsmérová
Študijný program: Informatika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: 9.2.1. informatika
Typ záverečnej práce: Bakalárska práca
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Behaviorálne aspekty vybraných častí ransomware

Názov EN: Behavioral aspects of selected parts of ransomware

Cieľ:
1) Analyzovať ransomware a jeho známe typy
2) Porovnať jednotlivé prístupy k analýze ransomware
3) Navrhnúť a porovnať metódy detekcie a obrany voči ransomware

Literatúra:
[1] OKTAVIANTO, Digit; MUHARDIANTO, Iqbal. Cuckoo Malware Analysis. Packt Publishing Ltd, 2013.
[2] SIKORSKI, Michael; HONIG, Andrew. Practical malware analysis: the hands-on guide to dissecting malicious software. no starch press, 2012.
[3] MARAK, Victor. Windows Malware Analysis Essentials. Packt Publishing Ltd, 2015.
[4] LIGH, Michael, et al. Malware analyst's cookbook and DVD: tools and techniques for fighting malicious code. Wiley Publishing, 2010.

Vedúci: RNDr. JUDr. Pavol Sokol, PhD.
Konzultant: Mgr. Ladislav Bačo
Ústav : ÚINF - Ústav informatiky
Riaditeľ ústavu: prof. RNDr. Viliam Geffert, DrSc.

Dátum schválenia: 09.05.2018

Poďakovanie

Ďakujem vedúcemu tejto práce RNDr. JUDr. Pavlovi Sokolovi, PhD. a konzultantovi Mgr. Ladislavovi Bačovi za nápady a odborné rady. Taktiež sa chcem poďakovať Bc. Jánovi Kotradymu za jeho pomoc a usmernenie v prípade nejasností.

Abstrakt

Jednou z top 15 hrozieb podľa ENISA Threat Landscape je ransomvér. Ransomvér môže mať obrovský negatívny dopad pre ľudí aj organizácie. Príkladom je použitie ransomvéru WannaCry, ktorý v máji 2017 spôsobil ohromenie systémov vo viac ako 150 krajinách a ktorý ohrozil zdravotnícke systémy v Anglicku alebo na Slovensku. V rámci tejto práce sa venujeme identifikácii rôznych typov ransomvéru pomocou ich behaviorálnych znakov (napr. spôsobu komunikácie, šifrovania súborov a pod.). Na získanie vlastností ransomvérov využívame systém na statickú a behaviorálnu analýzu malvéru – Cuckoo Sandbox. Následne sme analyzovali výsledky zo systému a pomocou triediaceho algoritmu ich delíme do tried. Súčasťou tejto práce je aj návod na konfiguráciu domáceho laboratória Cuckoo systému.

Kľúčové slová: *ransomvér, malvér, Cuckoo Sandbox, behaviorálna analýza, triedenie.*

Abstract

One of the top 15 threats by ENISA Threat Landscape is ransomware. It can have a huge negative impact on people and organizations. An example is WannaCry ransomware, which in May 2017 caused system paralysis in more than 150 countries and threatend healthcare systems in England or Slovakia. This work's focus is on identification of different types of ransomwares by analyzing their behavioral features (e.g., how they communicate, how they encrypt files, etc.). To get ransomware's features we are using system for static and behavioral analysis of malware - Cuckoo Sandbox. After that we are analyzing the reports from this system and by the help of clustering algorithm we devide rasomwares into clusters. This work also includes instructions for configuring the domestic laboratory of Cuckoo system.

Keywords: *ransomware, malware, Cuckoo Sandbox, behavioral analysis, clustering.*

Obsah

Úvod

V posledných rokoch sa už bežne stretávame s pojmom ransomvér. Dôvodom je to, že ransomvér útokov je stále viac a viac. Tento fakt potvrdzuje aj počet nových rodín. Podľa jedného výskumu bolo 400-krát viac rodín v roku 2016 ako v roku 2015 a na rok 2017 mali podobné predpoklady. Ďalším výhražným faktorom je aj WannaCry útok z mája 2017, kedy boli napadnuté Windows systémy vo viac ako 150 krajinách a týmto útokom ohrozili napríklad aj zdravotnícke systémy v Anglicku.

V tejto práci sa zameriavame na identifikáciu rôznych vlastností ransomvérov. Ďalším cieľom je porovnanie metód používaných pri analýze ransomvérov. Súčasťou práce je aj návod na konfiguráciu domáceho laboratória na analýzu ransomvéru pomocou Cuckoo systému.

V prvej kapitole sa hlbšie venujeme ransomvéru, jeho vlastnostiam, jednotlivými krokom, ktoré vykonáva. V tejto kapitole sme popísali aj to ako sa vyvíjal ransomvér počas rokov. Pre statickú a dynamickú analýzu ransomvérov existujú rôzne metódy, ktorých pozitíva a negatíva sme načrtli v druhej kapitole. V nasledujúcej kapitole sme priblížili rôzne online a desktop nástroje pre analýzu ransomvérov. Keďže v tejto práci sme sa rozhodli analyzovať ransomvér pomocou dynamickej analýzy, vybrali sme si jeden z dostupných nástroj, pomocou ktorého sme to mohli uskutočniť. Pri inštalácií tohto nástroja, Cuckoo Sandbox-u, sa môžu vyskytnúť rôzne problémy a preto vo štvrtej kapitole sme vytvorili jednoduchý návod. Taktiež sme uviedli niekoľko nastavení a opatrení, ktoré sú dôležité pri analýze ransomvérov. V poslednej kapitole sme zanalyzovali výstup tohto systému a na základe výsledkov sme vzorky rozdelili do tried.

1 Ransomvér

Ransomvér útokov je každým dňom viac a viac. Takýto rýchly rozvoj zapríčiňuje aj to, že ľudia alebo aj organizácie nemajú vhodné nástroje na to, aby takýmto útokom zabránili. Ransomvér sa tak dostane k citlivým údajom, ktoré ak obeť chce dostať naspäť musí zaplatiť istú sumu. Tým ale bohužiaľ podporujú vývoj nových, zákernejších rodín ransomvérov. Preto je dôležité, aby sme nenechali ransomvér útočníkom dosiahnuť svoj cieľ a teda aby sme útok odhalili čím skôr.

1.1 Definícia malvéru

Malvér (malware, „malicious software“) je zastrešujúci pojem pre viacero druhov škodlivých programov. Akýkoľvek program, ktorý nejakým spôsobom ublíži používateľovi, počítaču alebo sieti, môže byť považovaný za malvér [20]. K malvérom zaraďujeme vírusy, červy, trójsky kôň, advér, ransomvér atď. Ich hlavnou črtou je šírenie a vykonávanie nejakých škodlivých aktivít, pričom pre každú z nich sú tieto vlastnosti špecifické. Tieto typy škodlivého programu môžu vniknúť do zariadení, poškodiť ich, monitorovať činnosti na Internete alebo vykonávajú iné škodlivé aktivity.

1.2 Definícia ransomvéru

Slovo **ransomvér (ransomware)** môžeme rozdeliť na 2 časti „ransom“ a „ware“, kde ransom znamená výkupné. Druhá časť slova, ware, naznačuje, že ide o typ malvéru. Z toho môžeme vyvodiť všeobecnú definíciu. Ransomvér predstavuje typ malvéru, ktorý žiada o výkupné. Výkupné môže žiadať za odblokovanie zariadenia alebo za dešifrovanie súborov. Podľa toho môžeme ransomvér deliť na dva typy [4]:

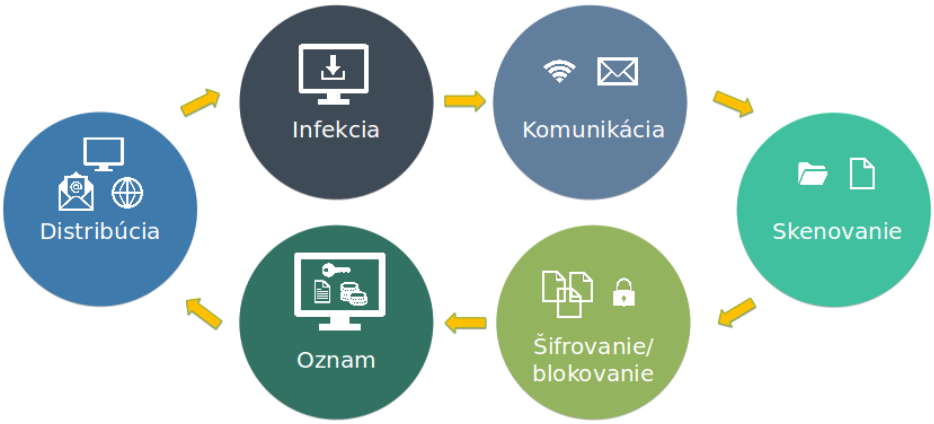
- šifrovací ransomvér a
- blokovací ransomvér.

Blokovací ransomvér zablokuje prístup k operačnému systému alebo softvérovému vybaveniu. Na druhej strane šifrovací ransomvér zašifruje vybrané dáta v za-

riadení. Útočníci blokovací ransomvér využívajú v menšej miere, keďže na odstránenie ransomvéru postačujú iba minimálne technické znalosti, kým pri šifrovacom ransomvéri je veľmi obťažné sa dostať k pôvodným súborom. Vo všeobecnosti oba typy ransomvéru sú úspešné aj z dôvodu, že útočník využíva nátlak na používateľov. Príkladom nátlaku na používateľa je možnosť zverejnenia alebo trvalého zmazania súborov. Používatelia následne zo strachu zaplatia výkupné. Zaplatenie peňazí avšak nezaručuje ešte vrátenie zariadenia/súborov do pôvodného stavu. Útočníci žiadajú výkupné vo väčšine príkladov v tzv. digitálnych menách. Príkladom je Bitcoin - kryptomena, pomocou ktorej môžu útočníci ostať v anonymite.

1.3 Spôsob fungovania ransomvéru

K tomu aby sa pomocou ransomvéru dosiahol cieľ, ktorý zamýšľa útočník, vykonáva malvér niekoľko typických krokov[1]. Tieto kroky sú zobrazené na obrázku 1.



Obr. 1: Typické kroky vykonávané ransomvérom.

Distribúcia - najčastejšie využíva štandardné metódy šírenia malvéru, ako sú napríklad e-mailové správy, ktoré obsahujú škodlivú prílohu. Ďalej využívajú zraniteľné webové sídla, po ktorých návšteve môže dôjsť k stiahnutiu ransomvéru do zariadenia. Často sa zneužíva aj zraniteľnosť iných aplikácií. Ransomvér sa zvykne šíriť aj pomocou botnetov.

Infekcia nastáva po tom, ako sa dostane ransomvér do zariadenia. V tej chvíli sa spustia procesy, ktoré potrebuje na vykonávanie škodlivých aktivít. V tomto kroku napríklad generuje jedinečný identifikátor pre napadnuté zariadenia. Súčasne vypne

[1]Dostupné na https://www.mcafee.com/kr/resources/white-papers/wp-understanding-ransomware-strategies-defeat.pdf

niektoré funkcie zariadenia, aby ostal neodhalený a aby mohol ďalej vykonávať svoju činnosť. Môže napríklad vymazať bod obnovy, čím zabráni obnoveniu zašifrovaných súborov.

Ďalšiu etapu predstavuje **komunikácia** ransomvéru s riadiacim serverom (tzv. C&C serverom), aby získal kľúče, ktoré sú potrebné na zašifrovanie údajov. Niekedy sa však stáva, že tento krok ransomvér nevykonáva, keďže útočník sa snaží minimalizovať sieťovú komunikáciu. V tomto prípade sú kľúče už súčasťou ransomvéru.

Po etape komunikácie prebieha etapa **skenovania**, pri ktorej ransomvér prehľadáva zariadenia a hľadá súbory, ktoré môžu byť dôležité pre používateľa a ktoré sú ťažko nahraditeľné. Pre tento účel hľadá najčastejšie používané súbory a súbory s príponami .jpg, .docx, .pptx, .pfd a pod. Blokovacie ransomvéry tento krok nevykonávajú.

Šifrovanie je ďalším a zároveň najhlavnejším bodom pri tomto útoku je zašifrovanie nájdených súborov z predošlého kroku. Okrem zašifrovania sú súbory často aj premenované a premiestnené. V tomto kroku blokovacie ransomvéry zablokujú prístup k operačnému systému alebo softvérovému vybaveniu.

V poslednej etape ransomvér **oznámi** používateľovi, že jeho zariadenie bolo napadnuté, koľko má zaplatiť (výšku výkupného) a kam je potrebné peniaze poslať. Môže to oznámiť v textovom dokumente, v okne vlastnej aplikácie, ako obrázok pozadia pracovnej plochy alebo spustením zvukovej nahrávky.

1.4 História ransomvéru

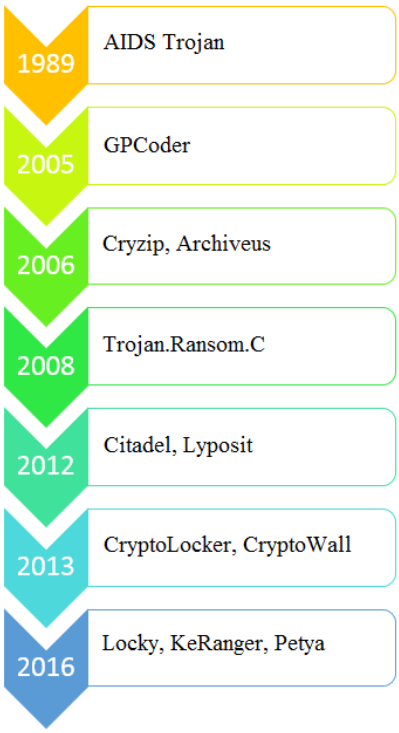
Prvý ransomvér sa objavil už v roku 1989 pod názvom **AIDS Trojan** (tiež známi ako PC Cyborg). Tento ransomvér bol vytvorený biológom Dr. Joseph Poppom [17]. Šíril sa pomocou floppy diskov, ktoré boli rozdané na medzinárodnej konferencii týkajúce sa choroby AIDS. Pôvodný autoexec.bat súbor bol nahradený iným, ktorý rátal počet reštartov počítača a po dosiahnutí limitu pomocou symetrickej kryptografie zašifrovali názvy súborov. Následne napadnutým osobám oznámili, aby poslali 189 dolárov vo forme šeku do Panamy. Tento ransomvér avšak nebol príliš úspešný, keďže v tom čase málo ľudí malo počítač, používal slabé šifrovacie techniky a zaplatenie výkupného bolo ťažko realizovateľné.

Po dlhých rokoch bez ransomvéru, resp. podobne sa správajúceho malvéru sa v roku 2005 začali šíriť ransomvéry, ktoré patrili do rodiny **PGCoder** (GP Code, Trojan.Gpcoder). Táto rodina využívala symetrické šifrovanie. Od roku 2006 sa začali šíriť internetom ďalšie rodiny ransomvéru. Ransomvér **Cryzip** prehľadával súbory s konkrétnymi príponami, zašifrované súbory presunul do heslom chráneného adresára a pôvodné súbory vymazal. Ďalší ransomvér **Archiveus** fungoval veľmi podobne s tým

rozdielom, že ako výkupné žiadali obete, aby nakúpili lieky z online stránky [16].

Trojan.Ransom.C sa objavil v roku 2008 a bol prvým ransomvérom, ktorý aj zablokoval počítač [16].

Od roku 2011 sa objavilo veľké množstvo nových typov ramsomvéru a začali sa objavovať rodiny ransomvérov, ktoré sa tvárili ako lokálne policajné orgány. V roku 2012 sa objavili nástroje na ľahšiu tvorbu a šírenie ransomvérov. Takýmto nástrojmi boli **Citadel** alebo **Lyposit**. Lypositom bola vytvorená aj rodina ransomvéru **Reveton**, ktorá oznámila správu, že na počítači sa vykonávala nejaká kriminálna aktivita a že počítač bol zablokovaný policajným orgánom. Výkupné mali zaplatiť vo forme kupónov [17, 21]



Obr. 2: Dôležité roky vo vývoji ransomvéru.

Jeden z najznámejších typov ransomvéru je **CryptoLocker**, ktorý sa objavil v roku 2013. Tento a ďalšie podobné typy nevyužívali len sociálne inžinierstvo, teda nielen vystrašili obete, ale využili aj ich slabé poznatky o počítačovej technike. Používali verejné a súkromné kľúče pre šifrovanie súborov. Zo začiatku bol šírený známym botnetom Zeus, neskôr pomocou mailov. Obete pôvodného CryptoLockeru mali 3 dni na zaplatenie. Po prekročení tohto limitu výška výkupného sa rapídne zvýšila. Obete mali možnosť platiť v dolároch alebo bitcoinoch. V nasledujúcich rokoch sa objavovali

ďalšie varianty CryptoLockeru, ktoré stále boli o niečo sofistikovanejšie a pribudol aj ransomvér **Cryptowall** a jeho rôzne verzie. Do roku 2015 útočníci využívajúci ransomvér vyzbierali 20 miliónov dolárov. Z tohto dôvodu verejné bezpečnostné orgány (napr. FBI) upozornili verejnosť a spoločnosti na túto hrozbu [17, 16, 21].

V roku 2016 sa začala šíriť e-mailová správa s Microsoft Office dokumentom v prílohe, ktorý obsahoval škodlivé makrá. V tomto roku boli prvýkrát napadnuté aj počítače s operačným systémom macOS [17, 16]. V tomto čase sa takisto prvýkrát vyskytol ransomvér **Petya**, ktorý nešifruje používateľské údaje, ale systémové súbory. Dokonca prepísal zavádzaciu tabuľku pre operačné systémy (Master Boot Record), čím znemožnil normálne fungovanie operačného systému [17, 16].

2 Analýza ransomvérov

Analýza ransomvérov je proces, pomocou ktorého sú skúmané vzorky ransomvérov, aby sme lepšie pochopili ako fungujú, aké škody môžu spôsobiť, ako sa dajú odhaliť a ako sa môžeme voči ním brániť. Keďže ransomvér je typ malvéru, pre analýzu ransomvérov exitujú rovnaké metódy ako pre malvéry. Preto v tejto a nasledujúcej kapitole ďalej spomíname metódy a nástroje pre analýzu malvérov s niekoľkými dodatočnými informáciami, ktoré platia špeciálne pre ranomsvéry.

2.1 Typy analýz malvéru

Pre analýzu malvérov existujú dva základné prístupy [20]:

- statická analýza a
- dynamická analýza.

Tieto dva typy analýzy ďalej môžeme deliť na základnú a pokročilú.

Každá z metód má svoje výhody a nevýhody. Výber analýzy závisí od výsledku, ktorý chceme dosiahnuť. Pre získanie lepších výsledkov je vhodné spojiť viaceré z metód, ktoré si detailnejšie priblížime v nasledujúcich podkapitolách.

2.2 Statická analýza

Statická analýza malvérov znamená analýzu kódu aplikácie pred jeho spusteným, aby sa zistilo, či je schopný vykonávať nejaké škodlivé aktivity. Ak sa nájde nejaká taká časť, tak sa zabráni spusteniu vykonávateľného súboru.

Jeden z typov **základnej statickej analýzy** používa porovnávanie digitálnych odtlačkov (hashov). Digitálne odtlačky súborov sa porovnávajú s odtlačkami známych škodlivých súborov. Táto analýza sa spolieha na obrovskú zbierku známych škodlivých kódov, ktoré sú stále aktualizované. V tom spočíva aj jedna z nevýhod tejto analýzy, čo predstavuje ich zdĺhavé porovnávanie. Ďalším negatívom je aj skutočnosť, že malými úpravami v kóde ransomvéru alebo malvéru dochádza k zmene digitálneho odtlačku.

Úplne postačí, ak dôjde napr. k odstráneniu komentárov, pridaniu riadku kódu a pod. S týmito zmenami sa škodlivosť kódu nezmení alebo len v malej miere. Systém, ktorý porovnáva digitálne odtlačky, túto zmenu neodhalí a bude považovať rovnaký škodlivý kód za rozdielny. Tieto metódy základnej statickej analýzy sú síce rýchle, ale nemusia odhaliť niektoré dôležité behaviorálne vlastnosti a sú veľmi neefektívne voči prepracovaným ransomvérom [12].

Inou možnosťou je použitie rozličných antivírusových programov na odhalenie škodlivosti. Je užitočné použiť viac antivírusových programov, keďže rôzne programy využívajú rôzne metódy na rozpoznanie škodlivých programov. Táto metóda je skôr len užitočný prvý krok pri analýze ransomvérov.

Ďalším typom základnej statickej analýzy je zbieranie informácií z reťazcov a funkcií. Pomocou týchto informácií vieme zistiť niektoré funkcionality danej vzorky ransomvéru. Získaná informácia nám môže napríklad povedať, aké knižnice ransomvér používa [20]. Taktiež môžeme vyhľadávať reťazce typické pre ransomvér, ktoré sa môžu nachádzať v súboroch alebo v názvoch vytvorených súborov. Takými charakteristickými slovami sú napríklad „ransom", „crypt" alebo „bitcoin".

Pokročilá statická analýza využíva reverzné inžinierstvo. Aplikuje sa hlavne pri prepracovaných ransomvéroch. Na takéto druhy ransomvérov, ako sme už spomínali, základná statická analýza nie je vždy vhodná. V niektorých prípadoch ani dynamická analýza, keďže sofistikované ransomvéry vedia odhaliť neštandardné (bezpečnostné) prostredie. V tomto prípade ransomvér prestane vykonávať svoju činnosť, a teda nevieme sledovať procesy, ktoré má vykonávať. Bezpečnostný analytik v tomto prípade spustí spustiteľný súbor v dissassembleri. Následne sleduje inštrukcie a pomocou tohto sledovania sa snaží odhaliť, čo daná vzorka ransomvéru robí. Tento typ statickej analýzy požaduje hlbšie poznatky týkajúce sa disassembleru, operačného systému atď. Tieto požiadavky môžeme považovať za nevýhodu tohto typu analýzy [20].

2.3 Dynamická analýza

Druhým prípadom analýzy malvérov je dynamická analýza. Detekcia pomocou **dynamickej/behaviorálnej analýzy** znamená pozorovanie vykonávaných procesov na určenie, či niektoré z nich majú škodlivé úmysly. Akýkoľvek proces so zlým správaním bude označený ako nebezpečný a ukončí sa [12].

Na to, aby sa mohli pozorovať procesy vykonávané malvérom, potrebujeme vytvoriť bezpečné prostredie. Takéto prostredie nazývame ináč aj ako „sandbox". Keď sa vytvorí prostredie, bezpečnostný analytik najmä pri analýze ransomvérov, sa nemusí báť, že stratí súbory, poškodí svoje zariadenie alebo ohrozí počítačovú sieť, v ktorej

sa zariadenia nachádza. Takéto prostredie môžeme vytvoriť na fyzickom alebo aj na virtuálnom stroji. Nevýhodou analýzy na fyzickom stroji je ťažké odstránenie ransomvéru po analýze. Naopak, na virtuálnom stroji to ide veľmi ľahko. Použitie fyzického stroja má výhodu v prípade, keď sa analyzuje taká vzorka ransomvéru, ktorá sa správa iným spôsobom vo virtuálnom stroji ako na fyzickom. Pri behaviorálnej analýze sa môžeme pozerať na rôzne aspekty, podľa ktorých môžeme špeciálne ransomvér sledovať. Môžeme napríklad sledovať operácie so súbormi, zisťovať šifrovanie súborov pozorovaním entropie alebo pozorovať premenovávanie súborov [6, 20].

Pri **pokročilej dynamickej analýze** je ransomvér spustený v ladiacom nástroji (debugger). Pomocou tohto nástroja môžeme sledovať detailne, aké inštrukcie dostáva ransomvér. Touto metódou môžeme tiež získať aj také údaje, ktoré nám neposkytne ani reverzné inžinierstvo. Môžeme napríklad vynútiť vykonávanie iných vetiev kódu zmenou v podmienkových krokoch v programe. Takto môžeme vidieť, ako sa ransomvér správa za iných okolností [6, 20].

3 Nástroje pre analýzu malvérov

Analýzu ransomvérov môžeme vykonávať pomocou rôznych nástrojov, ktoré využívajú rôzne metódy alebo kombináciu metód na odhalenie škodlivých aktivít. Analyzovať môžeme online alebo v počítači, či už pomocou nejakej aplikácie alebo priamym spúšťaním. Online nástroje poskytnú základné informácie o malvéri a dobrý prehľad o tom ako sa správa, ale väčšinou majú nejaké obmedzenia

3.1 Analýza malvérov pomocou online sandboxov

Jedným z najznámejších online nástrojov je **VirusTotal** [5], ktorý patrí pod spoločnosť Google. Môžeme nahrať alebo poslať mailom rôzne súbory do veľkosti 256MB ako napríklad obrázky, Android aplikácie, Windows spustiteľné súbory, JavaScript kódy alebo aj URL adresy. Výsledok obsahuje informácie o tom, či boli odhalené podporovanými antivírusovými programami, k akým súborom pristúpil daný malvér, akým spôsobom pracoval s registrami a ako sa správal. Tento nástroj síce poskytuje detailný popis daného malvéru, ale jeho nevýhodou je to, že výsledky ani vzorky sa nedajú stiahnuť.

Malwr [1] je spravovaný skupinou dobrovoľníkmi zo sféry bezpečnostných profesionálov. Je postavený na Cuckoo Sandbox-e a preto často sa používa, keď si niekto nemôže vytvoriť vlastný Cuckoo Sandbox. Môžeme pomocou nej analyzovať Windows spustiteľné súbory a rôzne dokumenty okrem URL adries a Android aplikácii. Niektoré vybrané vzorky sa dajú stiahnuť. Užitočným doplnkom Malwr sandboxu je MalwareViz [2], čo slúži na vizualizáciu výsledkov.

Hybrid Analysis [3] na backend-e používa Falcon Sandbox. Môžeme analyzovať rôzne súbory, napríklad PE, Office, PDF, APK, EML súbory a iné. Maximálna veľkosť nahrávaného súboru je 100MB. Pred spustením analýzy môžeme nastaviť niektoré nastavenia týkajúce sa dĺžky alebo prostredia v akom sa má vykonávať. Analyzuje a kombinuje dynamické a statické údaje s cieľom obohatiť výsledky analýzy napr.

[1] Dostupné na https://malwr.com/

[2] Dostupné na https://www.malwareviz.com/

[3] Dostupné na https://www.hybrid-analysis.com/

sledovaním inštrukcií alebo toku údajov. Taktiež ukladá obraz pamäte, aby mohol vykonať hlboký prehľad statickej analýzy vo fáze generovania správ. Pomocou tohto nástroja môžeme prezerať aj na už existujúce výsledky analýz.

Ďalším online nástrojom je **Metadefender** [14] od spoločnosti OPSWAT. Poskytuje komplexnú kontrolu obsahu súboru s najbežnejšími koncovkami, kontroluje softvér pre známe zraniteľnosti, poskytuje informácie o analýze od viac ako 40 antimalvérových programov a porovnáva IP adresy so známymi IP adresami, ktoré patria botnetom alebo iným hrozbám. Taktiež existuje rozšírenie pre prehliadač Chrome, ktoré prekontroluje každý stiahnutý súbor. Výhodou tohto nástroja oproti VirusTotal alebo Hybrid Analysis je to, že ponúka súkromné skenovanie, aby citlivé údaje ostali v bezpečnosti.

Android aplikácie sa čoraz častejšie stávajú obeťou malvérov. Na analýzu podozrivých Android aplikácií exituje online nástroj **AndroTotal** [11].

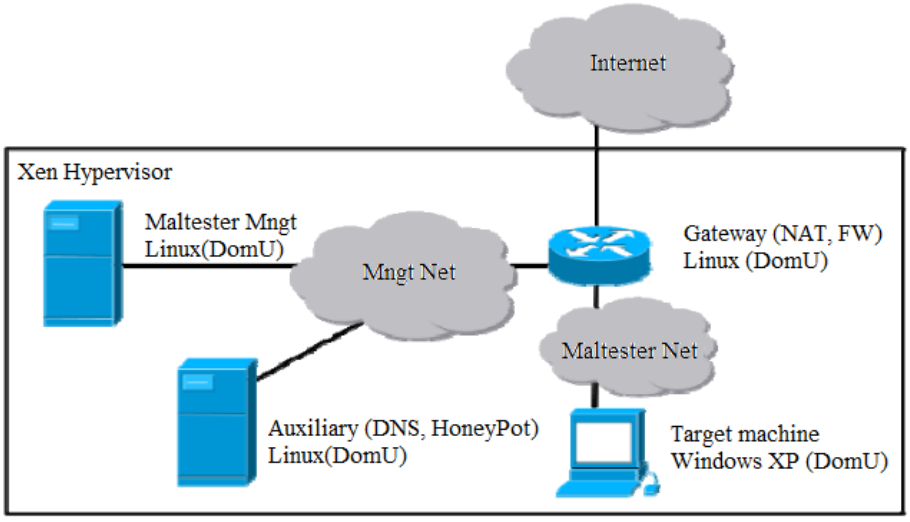
Spoločnosti Anubis a Wapwet, ktoré sa venovali online analýze malvérov, sa spojili a vytvorili firmu Lastline [10]. Naďalej sa venujú kybernetickej bezpečnosti,ale neposkytujú funkcionalitu online analýzy malvérov.

Síce tieto online nástroje sú dobré na získanie základných údajov, navyše pomocou niektorých nástrojov môžeme dostať aj detailnejšie informácií o analyzovanej vzorke, ale často môžeme naraziť na nejaké obmedzenia. Takými sú napríklad veľkosť súboru, typy analyzovaných súborov, možnosť spracovania získaných údajov a iné.

3.2 Analýza malvérov na fyzickom stroji

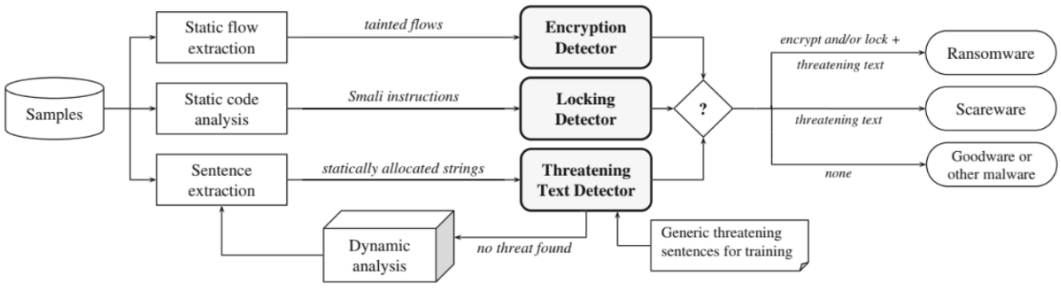
Jedným z nástrojov, ktorým môžeme vykonávať analýzu malvérov na fyzickom stroji je napríklad **Maltester**, ktorý použil Cabaj a spol. [3] V práci využívali metódy statickej analýzy na odhalenie vykonávaných krokov v infikovanom počítači v kombinácií s Maltesterom. Tento nástroj použili na dynamickú analýzu ransomvérov, na skúmanie sieťovej aktivity. Maltester využíva voľne dostupný Xen hypervizor, ktorý sa skladá zo štyroch virtuálnych strojov. Tieto nástroje sú súčasťou jedného fyzického serveru. Beží na Linux Debian operačných systémoch. Keď používateľ pridá malvér do systému vytvorí sa inštancia cieľového stroja. Tento stroj má softvér, ktorý vie komunikovať s riadiacim hostom a tak pridaná vzorka je automaticky posunutá a vykonaná v cieľovom systéme. Kvôli bezpečnostným opatreniam, cieľový systém nemá priamy prístup do internetu, ale dva virtuálne mosty sú pridané. Jeden slúži pre riadiace účely a druhý odchádzajúce pakety od cieľového stroja, keďže tie pakety môžu byť škodlivé. Navyše celá komunikácia je presmerovaná špeciálne nakonfigurovanou bránou. Táto brána sa stará aj o ukladanie celej sieťovej komunikácie počas analyzovania malvéru.

Po istom čase analýza je zastavená hostom, ktorý potom spustí dodatočné analýzy. Keď vzorka je analyzovaná prvýkrát, tak väčšinou komunikuje s reálnymi Command and Control (C&C) servermi, ale po čase je vytvorený emulátor, aby analýza danej vzorky nebola odhalená útočníkmi. Emulátor je vytvorený po pochopení protokolu používaného medzi malvérom a serverom a beží na dodatočnom stroji.



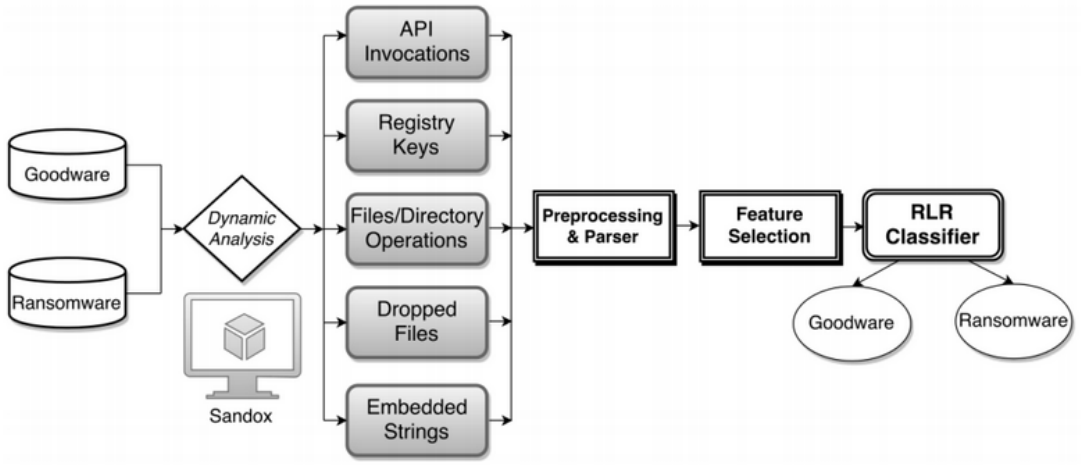
Obr. 3: Architektúra prostredia Maltester [3].

Andronio a spol. [2] skúmali Android ransomvéry. Pre tento účel použili plne automatizovaný nástroj zvaný **HELDROID**. Pomocou tohto nástroja mohli rýchlo a efektívne rozlíšiť, či už známe alebo neznáme ransomvéry od neškodných softvérov. Týmto nástrojom vedeli odhaliť aj to, či ransomvér chce šifrovať dáta alebo zablokovať zariadenie. V tomto nástroji využívajú aj statickú aj dynamickú analýzu. Statickú analýzu využívali na získanie tých postupnosti volaní funkcií, ktoré by mohli naznačovať snahu o zablokovanie alebo zašifrovanie zariadenia. Ďalej pomocou statickej analýzy vyhľadávali výhražné texty. Dynamická analýza sa použije iba v prípade, ak pomocou statickej analýzy žiadne výhražné texty neboli nájdené. Výhodou tohto nástroja je, že nie je špeciálne iba pre niektoré typy rodín a tým pádom dokáže odhaliť aj nové rodiny ransomvérov. Nevýhodou je, že sa sústreďuje iba na ransomvéry pre mobilné zariadenia a jeho prerobenie pre iné zariadenia je netriviálne. Pri testovaní nástroja ukázali, že z ich dataset-u HELDROID dokáže odhaliť všetky ransomvéry. Dataset okrem známych a neznámych ransomvérov obsahovali aj neškodné softvéry a malvéry. Síce sa vyskytli aj falošne pozitívné výsledky, ale bolo ich zanedbateľne málo.



Obr. 4: Architektúra systému HELDROID [2].

Strojové učenie pre dynamickú analýzu a triedenie ransomvérov využili Sgandurra a spol. [19] v nástroji **EldeRan**. Ich cieľom bolo zistiť, či sa dajú ransomvéry odhaliť ešte pred nakazením počítača pomocou niekoľko charakteristických vlastností ransomvérov. Ako prvé EldeRan dynamicky analyzuje v Cuckoo Sandbox-e ransomvéry a iné neškodlivé softvéry. Po tejto analýze sa vyberú kľúčové vlastnosti pomocou algoritmu. Ako posledné matica týchto vlastností je použitá na triedenie vzoriek na ransomvéry a neškodné aplikácie. Ukázali, že kombináciou týchto troch krokov EldeRan dokáže s 96,3-percentnou úspešnosťou odhaliť ransomvéry v prvotných fázach. Navyše vie odhaliť aj nové rodiny ransomvérov. Neodhalené ostali iba tie, ktoré dlhšiu dobu nevykonávajú žiadne kroky v počítači alebo tie, ktoré potrebujú interakciu používateľa pre vykonávanie ďalších škodlivých aktivít.



Obr. 5: Architektúra nástroja EldeRan [19].

Kharaz a spol. [9] predstavili nástroj **UNVEIL**, ktorý využíva dynamickú analýzu. Tento nástroj pri prvom kroku na dynamickú analýzu používa systém Cuckoo Sandbox. V tejto fáze sa automaticky generuje nové používateľské prostredie a súbory.

Autori dbali na to, aby tieto vygenerované súbory boli reálne, s rôznymi cestami a s rôznymi časovými pečiatkami. Následne na odhalenie šifrovacích ransomvérov monitoruje aktivitu súborového systému. Na odhalenie blokovacieho ransomvéru screenshoty sú vytvorené, v ktorých sa vyhľadávajú typické reťazce. Autori ukázali, že pomocou nástroja UNVEIL dokázali identifikovať väčšinu ransomvérov. Nevýhodou tohto nástroja je, že neodhalí tie ransomvéry, ktoré dokážu identifikovať automaticky vygenerované používateľské prostredie. Neidentifikované môžu ostať aj tie, ktoré nešifrujú celé súbory a tak aktivita súborového systému nie je nápadne veľká.

	Statická analýza	Dynamická analýza	Strojové učenie
Maltester	X	X	
HELDROID	X	X	
ElderRan		X	X
UNVEIL		X	

Obr. 6: Porovnanie metód jednotlivých nástrojov.

V popisoch nástrojov ste si mohli všimnúť, že väčšinou sa využívajú kombinácie metód na získanie čo najlepších výsledkov. Taktiež môžeme vidieť, že každá z nástrojov používa dynamickú analýzu. Tento faktor nám naznačuje ako užitočná a efektívna je táto metóda. Dynamická analýza je dobrou voľbou, keď sa pozeráme na vzťah medzi náročnosťou použitia a kvalitou výsledkov.

4 Cuckoo Sandbox

Ako sme vyššie spomenuli, vykonávanie behaviorálnej analýzy vyžaduje použitie bezpečného prostredia, ktoré nám tento typ analýzy umožní. Príkladom je Cuckoo Sandbox [7], čo je voľne dostupný automatizovaný bezpečnostný systém na analýzu rôznych typov škodlivých súborov v bezpečnom odizolovanom prostredí. Cuckoo Sandbox začal ako projekt na Google Summer of Code v roku 2010. Bol vytvorený v rámci Honeynet Projektu a beta verzia bola zverejnená začiatkom roka 2011.

Medzi súbory, ktorý je možné analyzovať v rámci tohto bezpečnostného systému zaraďujeme:

- súbory spustiteľné vo Windowse,
- DLL súbory,
- PDF súbory,
- Microsoft Office dokumenty,
- webové stránky a iné.

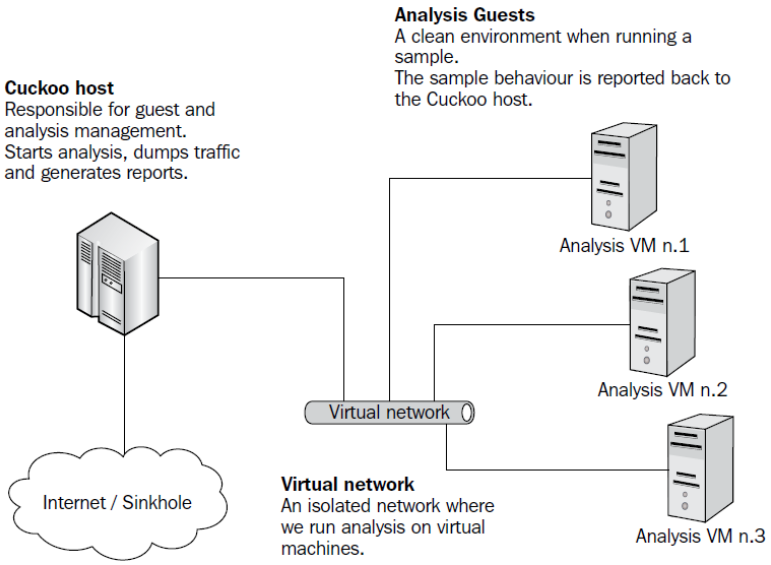
Pomocou tohto nástroja tieto súbory môžeme spustiť a sledovať ich správanie v reálnom čase. V priebehu niekoľko minút detailne vyhodnotí ich fungovanie. Z výstupu môžeme dostať rôzne informácie:

- zoznam volaní WIN32 API všetkých spustiteľných procesov počas analýzy,
- vytvorené, zmazané a stiahnuté súbory,
- časť pamäte, ktorá obsahujú procesy spustené malvérom,
- sieťovú komunikáciu,
- screenshoty obrazovky počas analyzovania a
- obraz pamäte celého systému.

Architektúra Cuckoo systému sa skladá z dvoch častí:

- **hostiteľského systému (host system)** - centrálny radiaci systém, ktorý má na starosti spustenie a analyzovanie malvérov.
- **hosťovaného systému (guest system)** – predstavuje systém, na ktorom sú spustené dané vzorky alebo súbory.

Po každej analýze sa hostiteľský systém postará o to, aby sa hosťujúci stroj vrátil do pôvodného stavu. Takto sa každá analýza vykoná v novom izolovanom prostredí.



Obr. 7: Ukážka prostredia Cuckoo systému [13].

V Cuckoo Sanbox-e hostiteľský systém je kompatibilný s viacerými operačnými systémami, avšak tvorcovia sanboxu odporúčajú GNU/Linux. Druhú časť tvorí jeden alebo viac virtuálnych, resp. fyzických strojov. Pre tieto stroje odporúčajú 64-bitový Windows XP alebo Windows 7 pre Windows analýzu, ale umožňujú aj použitie Mac OS X Yosemite pre Mac OS X analýzu a Debian pre Linux analýzu.

V našom riešení na fyzický stroj sme nainštalovali 64-bitový Linux Mint operačný systém. Pre hosťujúce systémy sme si zvolili Windows 7, keďže väčšina ransomvéru je určená pre tento operačný systém. Ako hypervizor sme si vybrali **VirtualBox** [15].

4.1 Inštalácia hostiteľského systému

Pred inštaláciou samotného Cuckoo systému je potrebné nainštalovať všetky balíčky, ktoré sú pre jeho funkciu nevyhnutné. Takými sú Python knižnice, keďže host komponenty pre Cuckoo sú písané v Pythone. Zatiaľ majú podporu iba pre Python 2.7. Pre použitie Django webového rozhrania je povinné mať nainštalované MongoDB.

Ďalej Cuckoo potrebuje virtualizačný softvér, z čoho rôzne existujúce podporujú, ale ako už bolo spomenuté, my používame VirtualBox. Ak chceme sledovať sieťovú aktivitu ransomvéru, sieťový analyzátor bude potrebný. Cuckoo štandardne si zvolil tcpdump. Pred vykonaním potrebných príkazov si môžeme vytvoriť používateľa špeciálne pre analýzu. Tohto používateľa potom treba pridať do skupiny „vboxusers".

```
$ sudo adduser cuckoo
$ sudo usermod -a -G vboxusers cuckoo
```

Spomínané potrebné balíčky môžeme nainštalovať nasledujúcimi príkazmi:

```
$ sudo apt-get install python python-pip python-dev python-setuptools libffi-dev
  libssl-dev libjpeg-dev zlib1g-dev swig mongodb tcpdump
$ sudo pip install m2crypto==0.24.0
```

Po týchto inštaláciách Cuckoo by už mal mať všetky potrebné balíčky, ale pred jeho inštaláciu ešte autori odporúčajú aktualizovať „pip" a „setuptools" knižnice. Cuckoo systém môžeme nainštalovať viacerými spôsobmi, napríklad zo súboru alebo klonovaním z ich oficiálneho úložiska na GitHub-e.

```
$ sudo pip install --upgrade pip setuptools
$ sudo pip install --upgrade cuckoo
```

4.2 Inštalácia hosťujúceho systému

Pod používateľom, ktorým chceme neskôr analyzovať si vytvorím jednu alebo viac virtuálnych strojov s operačným systémom Windows 7. Pre bezproblémový beh Cuckoo systému na hosťujúcich virtuálnych strojoch je tiež potrebné nainštalovať Python, taktiež verzie 2.7. **Python Pillow** [1] nie je povinné mať nainštalované, avšak odporúča sa, ak chceme mať prístup ku všetkým dostupným funkciám. Používa sa to na vytváranie screenshotov obrazovky počas analýzy. Ak sa rozhodneme využiť túto možnosť, je potrebné nainštalovať vhodnú verziu vzhľadom na verziu nainštalovaného Pythonu. Tvorcovia systému odporúčajú mať na hosťujúcom systému nainštalované aj dodatočné nástroje ako napr. webové prehľadávače, PDF prehliadač, kancelársky balík a iné.

4.3 Nastavenia hosťujúceho systému

Kvôli ľahšej práci pri pripravovaní systému najprv vykonáme potrebné nastavenia pre hosťujúci systém.

Jedno z dôležitých nastavení na hosťujúcich systémoch (virtuálnych strojoch) je vypnutie Windows Firewall a automatických aktualizácií. Aktualizácie je potrebné

vypnúť aj pre dodatočné softvéry, ktoré sme si nainštalovali. Ak sme si vybrali Windows 7, tak musíme vypnúť aj kontrolou používateľských kont (user account control – **UAC**).

Odporúčanie tvorcov systému v prípade sieťových nastavení je použiť **Host-Only Networking** štruktúru. Pre každý virtuálny stroj je potrebné nastaviť statickú IP adresu, keďže Cuckoo systém neobsahuje systém pre prideľovanie dynamických IP adries (DHCP).

Spravovanie komunikácie a výmenu dát medzi hostiteľským systémom a virtuálnymi strojmi má na starosti agent (skript napísaný v jazyku Python). Keďže tento agent je potrebné spúšťať na virtuálnom stroji, treba ho prekopírovať z pracovného adresára Cuckoo systému na virtuálny stroj. Ak chceme aby tento agent sa automaticky zapol pri každom zapnutí virtuálneho stroja, tohto agenta môžeme umiestniť do adresára „Pri spustení“ (Startup folder). Koncovku tohto súboru môžeme zmeniť z „.py” na „.pyw”, ak chceme vytvárať screenshoty a nechceme, aby sa nám tento agent zobrazoval na obrazovke.

Pri analýze ransomvérov je dôležité umiestniť do virtuálnych strojov veľký počet súborov rôznych typov. My sme umiestnili do našich virtuálnych strojov dokumenty, obrázky, videá, filmy, pesničky, prezentácie a tabuľky. Tieto súbory boli umiestnené na rôznych miestach. My sme ich umiestnili niekoľko na plochu, do priečinku Obrázky, Videá a Hudba a väčšinu do priečinka Dokumenty, kde sme vytvorili ďalšie rôzne adresáre. Ak máme všetko pripravené treba ešte vytvoriť snapshot, vypnúť a obnoviť virtuálny stroj.

```
$ VBoxManage snapshot "<Name of VM>" take "<Name of snapshot>" --pause
$ VBoxManage controlvm "<Name of VM>" poweroff
$ VBoxManage snapshot "<Name of VM>" restorecurrent
```

Po týchto nastaveniach môžeme jednoducho vyklonovať toľko virtuálnych strojov, s koľkými budeme chcieť pracovať. My sme mali dva virtuálne stroje, ktoré sme si nazvali cukoo1 a cuckoo2.

4.4 Nastavenia hostiteľského systému

Ako prvé v sieťových nastaveniach pre virtuálny stroj sme nastavili **Host-Only Network** podľa odporúčania autorov. Využíva sa vtedy, keď potrebujeme vytvoriť izolovanú virtuálnu sieť. Virtuálne stroje a hostiteľský systém vedia medzi sebou komunikovať, ale virtuálne stroje nevedia s nikým iným komunikovať mimo tohto virtuálneho prostredia. Takéto sieťové nastavenie je výhodné pri analýze malvérov, najmä pre analýzu ransomvérov. Dôležité je zabezpečiť, aby sa ransomvér nemal možnosť do-

stať sa k žiadnym súborom mimo virtuálneho stroja, keďže ich následne zašifruje. Ďalším krokom pri nastavení hostiteľského systému je nakonfigurovanie pracovného adresára CWD (Cuckoo Working Directory). Všetky konfigurovateľné komponenty, generované dáta a výsledky sú uložené v tomto adresári. Tieto nastavenia neskôr uľahčujú aktualizáciu na novšiu verziu. Ďalej je potrebné nastaviť konfiguračné súbory podľa vlastných potrieb. V rámci Cuckoo systému sa využívajú najmä tieto konfiguračné súbory [13]:

- cuckoo.conf,
- auxiliary.conf,
- <machinery>.conf,
- processing.conf,
- reporting.conf.

V konfiguračnom súbore **cuckoo.conf** môžeme nájsť základné nastavenia systému, informácie o všeobecnom správaní a o možnostiach analýzy. V tomto súboru musíme nastaviť používaný virtualizačný nástroj (machinery =) a v sekcii [resultserver] je potrebné nastaviť IP adresu nášho hostiteľského systému (ip =). V sekcii [timeouts] si môžeme nastaviť dĺžku trvania analýzy a čas, za ktorý sa musí vypnúť virtuálny stroj. My sme si nastavili tieto časy nasledovne: default = 30, critical = 10, vm_state = 15. Tieto hodnoty sme si zvolili tak, aby analýza netrvala príliš dlho, ale na druhej strane aby sme dostali dostatočné množstvo zaujímavých dát o činnostiach ransomvéru. Čas určený v „critical" je pridaný k hodnote zadaného do premennej „default". Keď sa tento čas prekročí, analýza sa považuje za neúspešnú.

Konfiguračný súbor **auxiliary.conf** popisuje a definuje možnosti pridávaných modulov, ktoré bežia súčasne s analýzou. Takýmto modulom je napríklad simulátor užívateľa, ktorý potvrdí vyskakovacie okná.

Konfiguračný súbor **<machinery>.conf** je označenie pre všetky hypervizory, ktoré systém podporuje. Používateľ opravuje konfiguračný súbor prislúchajúci k hypervizoru, ktrorý si stiahol. Keďže pre naše potreby sme si zvolili prostredie VirtualBoxu, tak názov konfiguračného súboru, ktorý potrebujeme upravovať bude virtualbox.conf. Obsahuje informácie o konfigurácií hosťujúceho systému. V tomto konfiguračnom súbore je potrebné nastaviť základné sieťové rozhranie, ktorá má používať (interface =) a názov jedného alebo viacerých čiarkou oddelených virtuálnych strojov (machines =). Následne pre každú z nich vyplniť označenie, platformu akú používa, IP adresu daného virtuálneho stroja a názov snapshot-u, ktorý má Cuckoo používať.

Prostredníctvom konfiguračného súboru **processing.conf** môžeme vypnúť/ zapnúť všetky moduly pre spracovanie dát. Pre naše účely sme nechali aktivované v tomto súbore nasledujúce sekcie: [analysisinfo] pre základné informácie, [behavior] pre získanie behaviorálnych vlastností, [network] na sledovanie sieťovej komunikácie a [snapshot] pre zaujímavosť, aby sme videli, či ransomvér zobrazí nejaký odkaz na ploche.

Konfiguračný súbor **reporting.conf** obsahuje informácie o generovaní výsledkov. V tomto konfiguračnom súbore stačí iba v sekcii [mongodb] zapnúť použitie mongo databázy (enabled =).

Po nastavení týchto konfiguračných súborov je potrebné ešte nastaviť smerovanie a smerovacie tabuľky tak, aby hosťujúce systémy (v našom prípade virtuálne stroje s operačným systémom Windows 7) mali prístup k s hostiteľskému systému (v našom prípade fyzické zariadenie s operačným systémom Linux Mint). Aby tieto nastavenia nezmizli po reštartovaní fyzického stroja, je potrebné vykonať patričné nastavenia.

5 Behaviorálne aspekty ransomvéru

5.1 Dataset

Pri analýze ransomvérov je dôležité s akými vzorkami sa pracuje. Čím je väčšia databáza vzoriek, o to lepšie výsledky môžeme dostať. Taktiež je dobré mať niekoľko vzoriek z každej analyzovanej rodiny.

Naš dataset bol od VirusTotal. Obsahoval vzorky, ktoré boli analyzované na tejto stránke v období 12. - 14.3.2018. Tieto vzorky už ale mohli byť aj skôr analyzované. Napríklad mali sme vzorky, ktoré boli prvýkrát analyzované na ich stránke v roku 2006. Dataset obsahoval rôzne Windows spustiteľné súbory a základné informácie o týchto spustiteľných súboroch v json formáte. Tieto súbory mali rovnaké názvy. Dostali sme ale aj také json súbory, ku ktorým v datasete neboli .exe súbory. To ukazujú aj počty. Dostali sme 64 313 Windows spustiteľných a 89 105 json súborov.

Pre naše účely sme vybrali iba niektoré .exe súbory. Vyhľadávali sme v json- och slovo „ransomware". Takýto reťazec sa nachádzal vo výsledkoch od antivírusových programov vtedy, keď antivírusový program pre dané správanie sa vzorky mal označenie ransomware. K týmto json súborom ak existoval spustiteľný súbor, tak sme ich prekopírovali do jedného priečinku pre ľahšiu prácu v ďalších krokoch. Tým sme získali 5227 vzoriek ransomvérov na analýzu.

5.2 Analýza pomocou Cuckoo systému

Pred samotnou analýzou ransomvérov ešte je dobré skontrolovať, či všetky miesta sú zablokované, odkiaľ by sa ransomvér mohol dostať do nášho fyzického stroja. Takým miestom môže byť napríklad zdieľaný priečinok, USB alebo zle nastavená sieťová konfigurácia. Taktiež pred spustením analýzy veľkého množstva vzoriek je dobré odskúšať na niekoľkých vzorkách, že pri nastaveniach, ktoré sme spravili aký veľký je výstup Cuckoo systému. Ak by nebolo dostatok miesta na fyzickom stroji pre dané nastavenia a počet vzoriek, je potrebné spraviť patričné kroky podľa vlastného zváženia a potrieb, aby počas analýzy nenastáli žiadne nečakané problémy.

Ak sme si tieto veci skontrolovali, môžeme spustiť Cuckoo Sandbox v termináli.

```
$ cuckoo
```

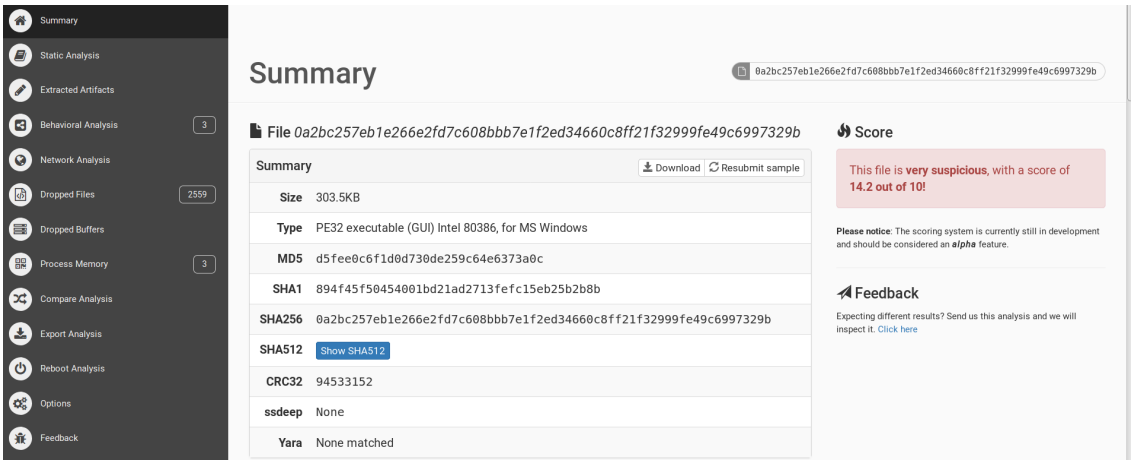
Vzorky pošleme do sandboxu v novom termináli pomocou príkazu

```
$ cuckoo submit /cesta/k/súboru,
```

kde je potrebné zadať adresár, obsahujúci naše pripravené vzorky. Pri tomto kroku každá vzorka dostane identifikačné číslo. Výsledky si môžeme pozrieť aj v ľubovoľnom webovom prehliadači. Pre rozbehnutie webového prehliadača je potrebné spustiť nasledujúci príkaz:

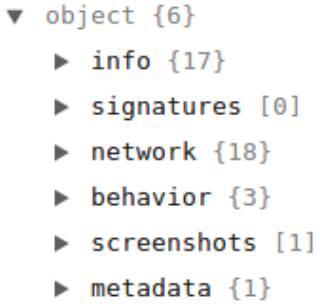
```
$ cuckoo web runserver
```

Po vykonaní by sme mali vidieť výsledky na adrese http://127.0.0.1:8000/.



Obr. 8: Ukážka Cuckoo webového rozhrania.

Zhrnutie výsledkov je možné aj stiahnuť, ale Cuckoo výsledky analýz ukladá aj na fyzický stroj. Nachádzajú sa v priečinku /home/<user>/.cuckoo/storage/analyses/, kde ďalej už sú rozdelené analýzy podla identifikátora, ktorý dostali. Tento priečinok sa naplní počas analýzy podľa toho, ako sme si nastavili konfiguračné súbory. Pre nás bol najdôležitejší súbor **report.json**, ktorého Štruktúra môžeme vidieť na obrázku 9.



Obr. 9: Štruktúra súboru report.json.

V kategórií „info" sa nachádzajú základné údaje ako napr.: začiatok a koniec analýzy, dĺžka trvania, typ súboru, identifikátor a iné. V „signatures" môžeme nájsť označenia pre kroky, ktoré ransomvér vykonával. Objavili sa nám tu napríklad údaje o tom, či sa ransomvér snažil detekovať, že sa nachádza v testovacom prostredí, či po vykonaní škodlivých aktivít sa snaží po sebe zmazať stopy, či vytvára koncovky súborov typické pre ransomvér atď. Pomocou týchto signatúr sme dostali približný náčrt o tom, ako sa vzorka správala v našom prostredí. V sekcii „network" je zhrnutie sieťovej komunikácie ransomvéru počas analýzy. O názvoch procesoch, o tom kedy alebo aké procesy čo vykonávali sa môžeme dočítať v časti „behavior". V posledných dvoch kategóriách sú umiestnené údaje o metadátach a o vytvorených screenshot-och.

5.3 Spracovanie údajov

Pre našu prácu boli najdôležitejšie sekcie týkajúce sa sieťovej komunikácie a behaviorálnych vlastností, aby sme získali prehľad o tom ako sa správa ransomvér. Snažili sme sa vybrať také údaje a v takom formáte, aby pri triedení sme dostali čo najlepšie výsledky.

Jednou z najcharakteristickejších vlastností ransomvérov je práca so súbormi. Aby sa táto vlastnosť zahrnula do triedenia, sledovali sme počty zvlášť vytvorených, zapísaných, otvorených, kopírovaných, presunutých a zmazaných súborov.

Ransomvér veľmi často, avšak nie stále, komunikuje na sieti s C&C servermi, aby získal napríklad kľúče potrebné na zašifrovanie súborov. Preto sme sledovali atribúty TCP, UDP a DNS. Ukladali sme len počty ich výskytov, keďže ransomvéry väčšinou používajú rôzne IP adresy. Ak by sme sledovali IP adresy, tieto údaje by nemali žiadnu účinnosť počas triedenia.

DLL (dynamic-link library) je skratka pre dynamicky spojenú knižnicu. Veľa funkcionalít operačného systému Windows sú poskytované týmito knižnicami. Obsahujú také inštrukcie, ktoré môžu používať rôzne programy. Z výsledkov analýz Cuckoo systému sa ukázalo, že ransomvéry používajú veľké množstvo rôznych DLL knižníc. Znova, aby sme dostali pri triedení najlepšie výsledky, základné knižnice sme vynechali a vybrali sme si menej zvyčajné, no často používané našimi vzorkami. Takými sú:

- ws2_32 - používa ju väčšina internetových aplikácií na spracovanie sieťových pripojení.
- wininet - obsahuje funkcie súvisiace s Internetom, ktoré používajú Windows aplikácie.
- dnsapi - používa sa DNS klientmi na overenie, či sa prekladá C&C doména

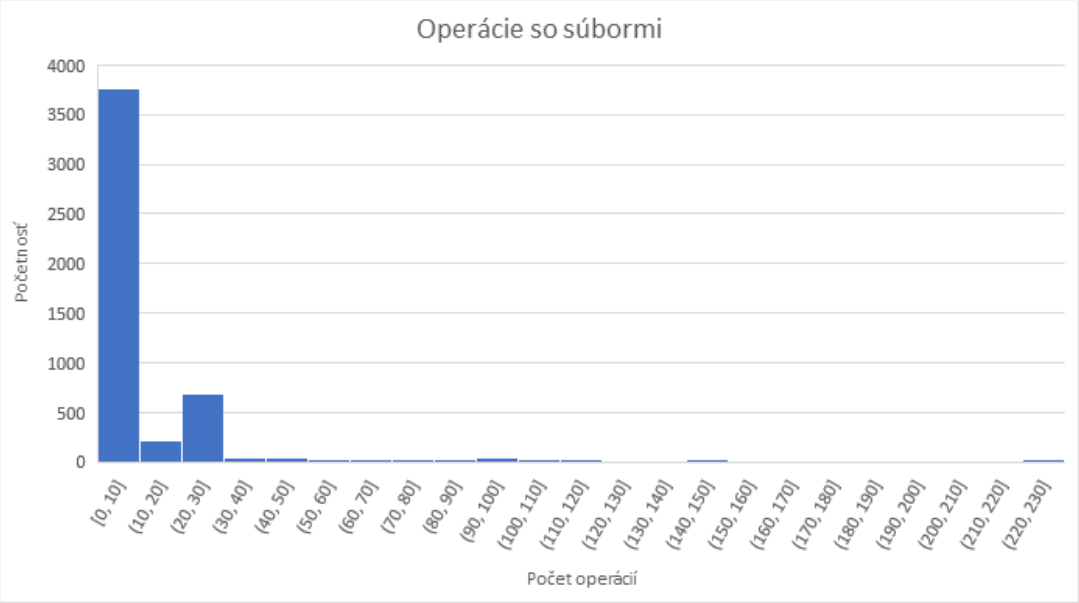
- netapi32 - obsahuje funkcie, pomocou ktorých sa zisťujú informácie o zdieľaných a sieťových diskoch
- urlmon - často sa používa na prácu s webovým prehliadačom
- wintrust - verifikuje dôveryhodnosť certifikátov
- cryptsp, crypt32 a cryptbase - obsahujú zaujímave kryptografické funkcie
- vboxmrxnp - detekcia Virtualbox - u
- secur32 - obsahuje bezpečnostné funkcie pre Windows

Vyskytli sa knižnice, ktoré aj keď obsahujú mnoho funkcií zaujímavých pre analýzu ransomvérov, nezahrnuli sme ich do zoznamu. Rozhodli sme sa tak preto, lebo obsahujú aj veľa iných základných funkcií, a je teda veľmi rozšírená. Takým je napríklad advapi32.dll, ktoré obsahuje kryptografické, ale aj základné funkcie na prácu s registrami. Následne už sme len vyhodnocovali, či daná vzorka použila knižnicu alebo nie.

5.4 Je to ransomvér?

V kapitole 2 bolo spomenuté, že antivírusové programy sú dobrým prvým krokom pri analýze malvérov alebo ransomvérov, ale nie úplne sa dá na nich spoľahnúť. Keďže my sme si vyberali vzorky z databázy na základe popisu vykonávaných krokov od antivírusových programov, v tejto fáze naše vzorky prešli druhým výberom. Ako kritérium sme si vybrali typickú vlastnosť ransomvérov a to prácu so súbormi. Pozerali sme sa na počty operácii a na ich početnosť. Hranicu sme určili na základe grafu, ktorý sa nachádza na obrázku 10. Ako môžeme vidieť veľa vzoriek pracovalo s menej ako 30 súbormi. Vzorku sme teda považovali za ransomvér, ak počet operácií so súbormi bol väčší ako 30. Dostali sme tak 808 vzoriek, ktoré sme považovali za ransomvér.

Po tomto výbere vzoriek ešte bolo potrebné upraviť atribúty, ktoré sme si vybrali na analýzu. Pri niektorých atribútoch bolo veľa rôznych hodnôt, ktoré by minimálne ovplyvnil triedenie. Preto hodnoty sme rozdelili do intervalov, kde ak i-ta hodnota padne do intervalu x, tak danú hodnotu nahradíme s x. Limity intervalov sme určili podobne ako pri druhotnom filtrovaní vzoriek. S tým nám vzniklo menej hodnôt pre vybrané atribúty.



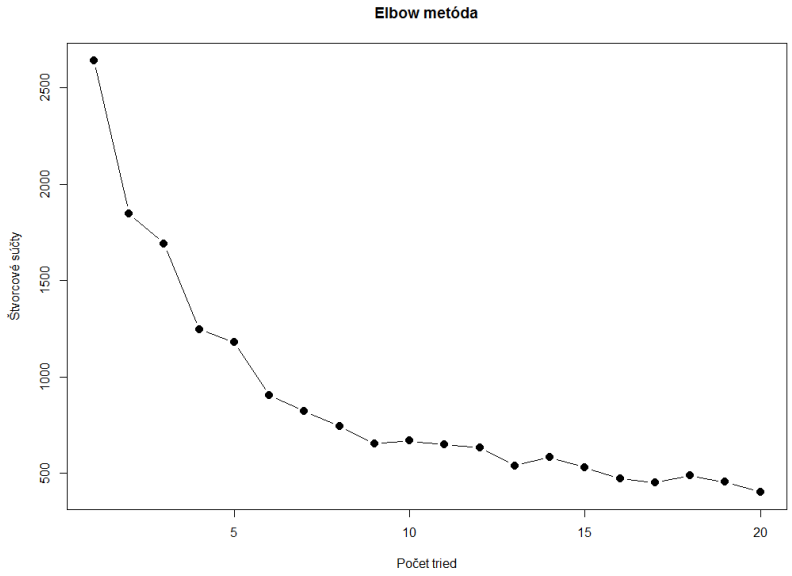
Obr. 10: Operácie so súbormi.

5.5 Triedenie

Triediaci algoritmus, ktorý sme si vybrali sa nazýva **K-mean**. Algoritmus rozdelí n vzoriek do K tried, kde každá trieda je reprezentovaná jednou centrálnou vzorkou. Tento prvok nazývame **centroid**. Prvky sa zaradia do tej triedy, ku ktorej centrálnemu prvku sú najbližšie. V každej triede centroid je taký prvok, ktorý má rovnaké atribúty ako vzorky. Jeho hodnoty sú priemery z hodnôt vzoriek, ktoré sa v triede nachádzajú. Centroid nemusí byť prvok z databázy. Hodnotu K je potrebné zadať.

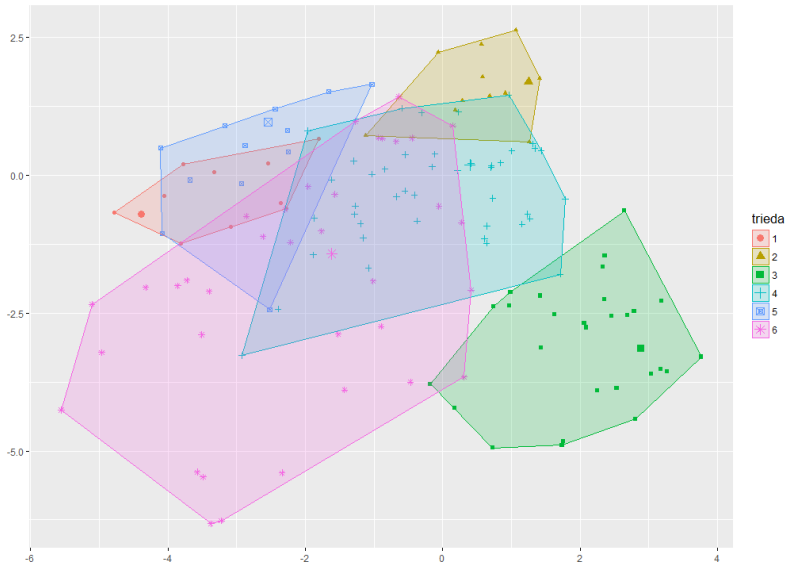
Na rozdelenie ransomvérov do tried sme použili jazyk R spolu s RStudio [18], kde K-means algoritmus už je implementovaný. Do funkcie stačí zadať dáta a počet tried. Ako výsledok dostaneme hodnoty jednotlivých centroidov a informáciu o tom, ktorá vzorka do ktorej triedy patrí.

To, aký je najvhodnejší počet tried, čo máme zadať ako parameter, môžeme zistiť viacerými spôsobmi. Jednou z nich náhodné vyberanie hodnôt a sledovanie, ktorou hodnotou dosiahneme najlepšie výsledky. Iná možnosť je použitie **metódy elbow**. Ideou tejto metódy je, že pustí K-means triedenie na vzorkách v rozmedzí hodnôt k (napr. pre k od 1 do 10) a pre každú hodnotu k vypočíta súčet kvadratických chýb (sum of squared errors - SSE). Tieto SSE hodnoty sa vykreslia pre každú hodnotu k do grafu. Vytvorí sa tak graf vyzerajúci ako rameno, kde lakeť ramena je hodnota K, ktorá je najvhodnejšia.



Obr. 11: Graf vytvorený elbow metódou.

Podľa výsledkov metódy sme sa rozhodli pre hodnotu K = 6. To znamená, že algoritmus K-means nám rozdelil ransomvéry do 6 tried. Na obrázku 12 sú znázornené vytvorené triedy, ktoré sme vykreslili pomocou knižnice factoextra [8].



Obr. 12: Triedy vytvorené K-means algoritmom.

5.6 Zhodnotenie

V súčasnej dobe sa rôzne nástroje na detekciu ransomvérov zameriavajú iba na niekoľko rodín. Sledujú vlastnosti, ktoré sú typické pre tieto rodiny. Nové rodiny ransomvérov, s novými vlastnosťami, každým rokom pribúdajú. Tento prístup nie je najvhodnejší, keďže nové vzorky ostanú neodhalené.

V práci sa ukázalo, že veľké množstvo rôznych ransomvérov sa dá rozdeliť do pár tried. Toto zistenie je užitočné pri vytváraní nástrojov, keďže je menej vlastností, ktoré je potrebné sledovať. Ak máme viac rodín s rozličnými vlastnosťami, tak nám narastá aj počet pozorovaných vlastností. Pri nájdení najmenšieho množstva tried, znížime aj počet sledovaných vlastností ransomvéru.

Záver

V práci sa venujeme problematike ransomvéru. Zameriavali sme sa na identifikáciu rôznych vlastností ransomvérov. Ďalším cieľom bolo porovnanie metód používané pri analýze ransomvérov. Súčasťou práce je aj návod na konfiguráciu domáceho laboratória na analýzu ransomvéru pomocou Cuckoo systému.

V prvej kapitole sme sa hlbšie venovali tomuto typu malvéru. Popísali sme rozdiely medzi jednotlivými typmi ransomvérov a uviedli dôležité roky vo vývoji ransomvéru. Typické kroky ransomvérov sme spomenuli a podrobne opísali tiež v tejto kapitole. Rôzne metódy existujú na statickú a dynamickú analýzu ransomvérov, ktorých pozitíva a negatíva sme načrtli v druhej kapitole. V tretej kapitole sa venujeme nástrojom, ktoré slúžia na analýzu malvérov a ransomvérov. Uviedli sme funkcie, výhody a nevýhody rôznych online a desktop aplikácií. Vo štvrtej kapitole sme popísali inštrukcie pre inštaláciu nástroja, ktorý sme si vybrali na analýzu. Systém, ktorý sme použili sa nazýva Cuckoo Sandbox. Upozornili sme aj na niekoľko nastavení, ktoré je potrebné skontrolovať špeciálne pred analýzou ransomvérov. Výstup z tohto systému sme následne pripravili na ďalšie spracovávanie. Pomocou týchto údajov a triediaceho algoritmu K-means sme dokázali rozdeliť viac ako 800 vzoriek ransomvéru do 6 tried.

Zoznam použitej literatúry

[1] AB, S. L. Python pillow.

[2] ANDRONIO, N., S. Z., AND MAGGI, F. Heldroid: Dissecting and detecting mobile ransomware. In *International Workshop on Recent Advances in Intrusion Detection* (2015), Springer.

[3] CABAJ, K., P. G. K. G., AND OSOJCA, D. Network activity analysis of cryptowall ransomware. *Przeglad Elektrotechniczny* (2015).

[4] CERT.BE. Ransomware whitepaper. Dostupné na internete: `https://www.cert.be/files/ransomware_whitepaper.pdf`, 2017.

[5] CHRONICLE. Virustotal.

[6] EGELE, M., SCHOLTE, T., KIRDA, E., AND KRUEGEL, C. A survey on automated dynamic malware-analysis techniques and tools. *ACM computing surveys (CSUR)* (2012).

[7] GUARNIERI, C., TANASI, A., BREMER, J., AND SCHLOESSER, M. Cuckoo sandbox: A malware analysis system. Dosupné z: `https://cuckoosandbox.org/`, 2012.

[8] KASSAMBARA, A., AND MUNDT, F. Cran - package factoextra.

[9] KHARRAZ, A., E. A. Unveil: A large-scale, automated approach to detecting ransomware. In *USENIX Security Symposium* (2016).

[10] LASTLINE, I. Lastline.

[11] MAGGI, F., VALDI, A., AND ZANERO, S. Andrototal: a flexible, scalable toolbox and service for testing mobile malware detectors. In *Proceedings of the Third ACM workshop on Security and privacy in smartphones & mobile devices* (2013), ACM.

[12] NIEUWENHUIZEN, D. A behavioural-based approach to ransomware detection. *Whitepaper. MWR Labs Whitepaper* (2017).

[13] OKTAVIANTO, D., AND MUHARDIANTO, I. *Cuckoo Malware Analysis*. Packt Publishing Ltd, 2013.

[14] OPSWAT. Metadefender.

[15] ORACLE, V. Virtualbox. Dosupné z: `https://www.virtualbox.org/`, 2015.

[16] RAJPU, T. S., E. A. Evolving threat agents: Ransomware and their variants. *International Journal of Computer Applications 164* (2017).

[17] RICHARDSON, R., AND NORTH, M. Ransomware: Evolution, mitigation and prevention. *International Management Review 13* (2017).

[18] RSTUDIO, I. Rstudio: A platform-independent ide for r and sweave.

[19] SGANDURRA, D., E. A. Automated dynamic analysis of ransomware: Benefits, limitations and use for detection. *arXiv preprint arXiv:1609.03020* (2016).

[20] SIKORSKI, M., AND HONIG, A. *Practical malware analysis: the hands-on guide to dissecting malicious software*. no starch press, 2012.

[21] ZAVARSKY, P., D. L. E. A. Experimental analysis of ransomware on windows and android platforms: Evolution and characterization. *Procedia Computer Science 94* (2016).

Prílohy

V prílohách sa nachádzajú nasledujúce položky:

Príloha A: Obsah CD

- Výstupné údaje v priečinku *Ransomware/Data*
- Použité skripty v priečinku *Ransomware/Scripts*
- Podrobnejší popis obsahu v *Readme.txt*